A STUDY OF THE EFFECTIVENESS OF A PHYSICAL SCIENCE INQUIRY COURSE

IN CHANGING THE ATTITUDES OF COLLEGE STUDENTS

TOWARD SCIENTIFIC METHODS

APPROVED BY SUPERVISORY COMMITTEE:

[redacted]

A STUDY OF THE EFFECTIVENESS OF A PHYSICAL SCIENCE

INQUIRY COURSE IN CHANGING THE ATTITUDES

OF COLLEGE STUDENTS TOWARD

SCIENTIFIC METHODS

by

JOHN HUGO MAULDIN, B.ENG.PHYS., M.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 1975

Abstract

A STUDY OF THE EFFECTIVENESS OF A PHYSICAL SCIENCE INQUIRY COURSE IN CHANGING THE ATTITUDES OF COLLEGE STUDENTS TOWARD SCIENTIFIC METHODS

The inquiry approach to teaching science is now being used at the college level, principally in physical science courses for the general student not majoring in science. Science courses for the general student can contribute to the level of scientific literacy in society, particularly if ability to use scientific inquiry is considered an important component of scientific literacy. For this study, it was assumed that favorable attitudes toward inquiry might lead to greater utilization of scientific inquiry by students as members of society.

In the inquiry method of teaching, the student is to learn from experiences with simple laboratory experiments. There are several approaches to the use of inquiry, but the most popular form involves the basic processes of 1) observing phenomena, 2) obtaining and comparing measurements on phenomena, 3) drawing conclusions or inferences from the observations, and 4) recognizing some of the assumptions underlying the conclusions. This form is called "guided discovery;" other forms might be more rigorous or more historical in approach. In the guided discovery method used in the physical science course on which this study was made, the students followed a prepared sequence of laboratory investigations, the results of which were left open-ended. Scientific methods of inquiry were used more as a learning method than as a rigorous approximation to science. The attitudinal effects of inquiry teaching might have been particular to the specific form of inquiry used. The inquiry course studied was not intentionally designed to change attitudes, yet some attitudinal effect is to be expected from any course or method of teaching.

The main question of the study was whether any attitudinal effects of inquiry teaching could be demonstrated. No known previous study of physical science courses, with or without inquiry, covered all three of these elements in this study: the science course was at the college level, the course used inquiry, and attitudes toward inquiry or scientific methods were investigated rather than general attitudes toward science. To avoid problems of finding another population with which to compare changes in attitude, an internal criterion for attitude change was used. The basic design of the study made use of the fact that the physical science course studied came as a two-semester sequence. The student decision whether or not to take the second semester of the inquiry course was taken as a behavioral manifestation of student attitude toward the first inquiry course. This unobtrusive measure of attitude was compared to attitude changes within the first course and to other possibly relevant variables. Multivariate linear regression analysis was used to find which variables predicted the decision to take a second inquiry course. Commonality analysis was included to separate the independent from the joint effects of the variables.

The variables of the study were treated in five sets:

1) Attitude level and change as measured with semantic differentials applied to three phrases about the inquiry aspect of the course.
2) Knowledge level and gain as measured with a 21-item multiple choice knowledge test on physical science and scientific methods.
3) General interest in science as measured with questions on the outside reading and scientific activities done by the student.

4) Background variables such as year, major, sex, GPA, and number of science courses already completed.

5) Instructor effect as measured with a semantic differential used to evaluate the instructor.

The sample consisted of 300 students on whom complete pretest and posttest data for the above variables was obtained. This sample was found to resemble the total population of students taking all 25 sections of Physical Science 303 at the University of Texas. Eighteen sections taught by 14 different instructors were included in the study. The sample was found to have less resemblance to the undergraduate population of university students not majoring in science or engineering. The students who continued from the first course to take the second course were identified by social security numbers (also used to match pretests with posttests).

The main result of the study was that attitude change in the first inquiry course was not found to be related to student decision to take the second inquiry course. The only variables found predictive of the decision were, in order of importance, a) the number of science courses the student had completed, b) the student's evaluation of the instructor independent of who the instructor was, and c) the student's level of knowledge but not the gain in knowledge. All of the variables together explained only 24.2% of the variance in the criterion (the decision to take the second course). The importance of the number of science courses completed seemed to be due to the fact that, on the average, the students needed two more science courses at the time they entered the first inquiry course. The second inquiry course was an obvious completion of the requirement unless there was reason to choose another science course.

Within the first inquiry course, the attitude change was significant and in the unfavorable direction. When this result was compared with the fact that 50% of the students showed a favorable attitude toward the first course by planning to take the second course, it seemed that the semantic differential measure of attitude toward inquiry did not coincide with the behavioral measure of attitude. The attitude shift in the first course cannot be confidently attributed to the effects of inquiry teaching. All variables together explained 50.3% of the variance in posttest attitude. The most important predictor of posttest attitude was again student evaluation of instructor.

The knowledge gain averaged two test items out of 21, and alpha reliability on the knowledge test was 0.68, low despite earlier efforts to improve the content of the test. Knowledge level or gain was not found related to attitude level or gain during the inquiry course.

There was indication of interaction among the semantic differentials because the correlation between posttest attitude toward inquiry and evaluation of instructor was high (0.53). Further studies should eliminate that problem as well as explore what student-instructor events might enter into student attitudes toward inquiry.

Dissertation study by John Hugo Mauldin
University of Texas at Austin, August 1975

Committee chairman: Robert N. Little
Committee members: Austin M. Gleeson
Addison E. Lee
John Roueche
Donald Veldman

TABLE OF CONTENTS

Chapter 1. Introduction to the Study

For more than a decade, various alternatives to the lecture-method of teaching science in college have been revived or developed, particularly for the general student not majoring in science or engineering. Educators have expressed a range of concerns and goals with respect to the improvement of science teaching. New forms of science courses, as well as conventional forms, can be assessed according to a general set of objectives for science teaching such as those stated recently by the National Assessment of Educational Progress Project (Education Commission of the States, 1972). Those objectives for a student in a science course are for the student to:

1. Know the fundamental aspects of science.
2. Understand and apply the fundamental aspects of science in a wide range of problem situations.
3. Appreciate the knowledge and processes of science, the consequences and limitations of science, and the personal and social relevance of science and technology in our society.

The third objective, having to do with appreciation and "relevance," was one inspiration for this present study. As stated in the Instructor's Manual (Little, 1973), these objectives form a planning basis for a physical science inquiry laboratory course in which this investigator taught for several years. That teaching experience showed that the effectiveness of the inquiry approach might in part be revealed by the attitudes of the students in the course.

Further motivation for this study of the attitudinal effects of inquiry teaching originated from an interest in the inquiry approach in general, as used in elementary, secondary, and college level science courses. Inquiry teaching is a method whereby the student directly experiences some phenomena in a laboratory atmosphere and is expected to attempt to infer concepts or laws from those observations. Ideally, the results obtained inductively from observation by the student might resemble concepts and laws already a part of the established body of scientific knowledge. The learning processes used are essentially scientific methods, and ideally in an inquiry course, all subject matter might be developed experientially with these methods. Lectures, textbooks, homework, and the like may be of secondary importance or may even be dispensed with. Henceforth, the term "inquiry course" will be used to refer to science courses in which the inquiry method of teaching is used.

Personal observation of the daily operation of an inquiry course for non-science college students showed a possible schism in the attitudinal effects of the course. On the one hand, the students seemed to enjoy being in the laboratory working with equipment, with each other, and with the instructor. On the other hand, evidence that they enjoyed carrying out simple scientific methods (such as observing and measuring phenomena accurately, comparing data, or generalizing from data) seemed difficult to establish objectively. Yet it might be important to investigate the effects of the course on attitudes toward the scientific methods that comprise inquiry. Under the assumption that a student might have to enjoy using inquiry in order that the student later retain some general ability at inquiry, the attitudinal effects of an inquiry course might contribute to a student's later attitude toward science and technology while a participating member of society.

<u>Societal effects of an inquiry course</u>

The majority of students do not major in science, but all do eventually participate in society. Some will teach in public schools where science is always a part of the curriculum. Therefore, the ultimate effect on society of the student's taking a course is particularly important when the course is one in science for the general student. At present, nearly all secondary students take a physical science course, which is taught in the inquiry mode wherever there are teachers trained or retrained to use inquiry. In Texas, for example, 86% of all ninth graders take a physical science course specifically designed to be taught by the inquiry approach (Texas Education Agency, 1972-73 Report). But it is estimated (Little, 1970) that in Texas only a few percent of the physical science teachers have received inquiry training. The retraining process is so slow that for some decades the most training that the majority of physical science teachers will have received will consist of their experiences in a ninth grade physical science course.

At the college level, many students encounter some form of science course requirements, usually minimal for those not majoring in science or engineering. Typically, these requirements have been met with general introductory, non-technical courses in the various sciences, often with some emphasis now on the societal problems that arise with that particular science. These courses are taught either by the lecture method or by lectures and laboratory exercises. An alternative approach is the laboratory-based inquiry approach, which is now being used in a number of physical science courses. Such courses are often begun to serve prospective elementary school teachers, but when they are opened to the remainder of the college or university population, they attract an increasing number of other majors needing to fill science requirements. Again, as an example, at the University of Texas at Austin, a physical science inquiry course was begun in 1971 for teachers. It now serves over 900 students in a two semester sequence, and teachers comprise only 17% of those who take it.

The inquiry approach to teaching science is not intended just as a way to encourage students to appreciate science in general. It is intended to expose the student in a first-hand manner to some of the scientific methods used in science, and thus it is in a position to foster appreciation of the actual processes by which scientists and others can acquire objective knowledge of the world. The distinction is made here between the general enterprise of science and the specific ways that science is done. There is need for both of these viewpoints in science teaching--the general and the specific, and this study is intended particularly to explore the specific. The inquiry course provides an ideal environment for this study because the main focus of the course is on specific scientific methods.

Insofar as taking or not taking science courses is a measure of general attitude toward science, the general student cannot be said to have a highly favorable attitude toward science. Moreover, some sciences are less favored than others. For example, in secondary schools in Texas, only 9% of the twelfth graders took physics in school year 1973-74 (Texas Education Agency Superintendent's report). Biology was more popular, with 47% of tenth and eleventh graders electing it. Those students interested in science tend to go to college, so that the typical proportion of students majoring in science and engineering is relatively large (26% of the undergraduates at the University of Texas). Aside from preferences for particular subject areas in science, it is possible that the methods of science as presented in various ways in these courses might affect those choosing to take science. Most of the courses are not inquiry courses, yet a study of the attitudes in an inquiry course might clarify to what extent attitudes toward scientific methods comprise general interest in and attitude toward science.

The third course objective from the Education Commission of the States' report stated that the student in the course is to "appreciate the knowledge and processes of science." The term "appreciate" raises the issue of whether a course is to be designed to change or affect student attitude. There are different views on this issue. Some critics would label attempts to change attitudes as "indoctrination" (Krathwohl, et al, 1964, p.18). Proponents of attitude change can take one of two views: either courses can be intentionally designed to change attitudes in a supposedly benevolent way (Wilkerson, 1972), or attitude change can be an inevitable and unintentional effect of a course. Subject matter such as science should be, or can be, depending on the view taken, appreciated as part of man's culture just as, for example, music or history (Krathwohl, et al, 1964).

An education research study is properly undertaken on the attitudinal effects of a course only after the purpose of any attitude changes in the course under study is clarified. This study was done with a college inquiry course elected by students to fulfill requirements. Although the course designer "hoped" that an appreciation of science and scientific methods would result for the student, no attempts were made to incorporate specific attitude-changing techniques in the course design. The fact that the course is voluntarily chosen by students who are partially informed about its laboratory approach is another reason that any attitudinal effects of the course are not of ethical concern.

It might be intrinsic to the inquiry method of learning that favorable attitudes toward scientific methods are developed. In the inquiry course, the student is intimately acquainted with the stages in the acquisition of objective knowledge. This personal participation in, and responsibility for, the development of the subject matter of the course might develop confidence and other attitudes favorably.

Attitudes toward scientific methods acquired in an inquiry course are a part of general attitude toward science, both during the course and later. As a member of society, a person's attitude toward science is a part of his/her "scientific literacy." The concept of scientific literacy includes not only knowledge of facts, laws, and methods, but also "the ability to use the fundamental aspects of science in everyday problem solving and in personal and public decision making" (Education Commission of the States, 1972). And it includes an appreciation of the scientific enterprise and of whatever facts, laws, concepts and inquiry skills the person has learned.

The public's need for scientific literacy, especially in regard to knowledge of science, was demonstrated by a study by Koelsche and Morgan (1964). They subscribed for six months (1962-63) to 22 of the largest general newspapers and 9 of the largest magazines. They clipped 3,000 science-related articles and found that the reader needed to know 175 different scientific principles and 693 scientific terms to read them. Since many of these articles related to public or business decisions about science, participation by the readers in any democratic decisions would involve not only their knowledge of the science involved but also their opinions and attitudes toward science. Their ability to use inquiry would also be pertinent because inquiry skills would be of assistance in assessing the facts as presented in the news media, in recognizing crucial assumptions, and in drawing conclusions on scientific issues. To apply inquiry to such issues would seem to presuppose a favorable attitude toward using inquiry.

In an article by Cohen (1964), the social context of science is discussed in relation to the promotion of scientific literacy through science teaching. Of his eleven goals for science teaching, the following are cited as being most relevant to scientific literacy for the general population (using his numbering):

2. To enable the citizen both to criticize and to appreciate the effects of the science on his society; to understand its history and its present alternative prospects.

3. To give a practical grasp of scientific methods of grappling with problems, at least sufficient for problems which the student will face in his individual and social life.

6. To understand the place of science among other intellectual and esthetic pursuits.

9. To provide our students with rich and various experiences of individual thinking and critical attitudes...and cooperative enterprise and mutual aid...

10. To help the pupil to understand the position of his future job or activity in the productive web of society...

11. To educate our students so that they may distinguish ends from means, probabilities from certainties, evidence from propaganda, questions from pseudo-questions, rational belief from superstition, and science from quackery.

Since science teaching goals such as the ones cited above seem to be presented in terms complimentary to the scientific enterprise, it would be worthwhile to present an opinion critical of science, particularly in reference to scientific literacy. As an example critical of science, Commoner (1963) has written that the proliferation of scientific knowledge in journals places more and more of science beyond the contemplation of the public. The implication is that students would have to learn more and more to keep up. [However, as Bruner points out (1962), inquiry teaching can be a way to give students tools to cope with excessive information by abstracting important patterns and looking for underlying unities.]

Commoner further claims that scientists are relying more and more on the "authority" of other scientists rather than remaining acquainted with how others developed their theories. This problem would seem to apply also to teaching non-science students, for the general student seems expected to take every fact and theory in science "on authority." [In an inquiry course, the student might be able to examine the origins of some facts and theories directly. Thus these criticisms might be taken as recommendations for inquiry teaching.]

A recent poll of the general public attitude toward science reveals an anomaly in the public's scientific literacy. The survey was done for the National Science Foundation in 1972 and reported in Science (26 Oct. 1973, p.369). It was found that the predominant emotion that the public had toward science and technology was "satisfaction and hope" by 49% and "excitement or wonder" by 23%. Six percent expressed "fear or alarm" and 6% expressed "indifference or lack of interest." In a list of nine prestigious professions, scientists ranked second, just below physicians. Fifty-four percent believed science and technology do more good than harm and 4% subscribed to the opposite view. Thirty percent believed that science will solve most major societal problems, and 47% that it will solve some. Simultaneously, 48% held science and technology responsible for some of society's problems, and 7% for most. When asked about priorities for tax monies, they put "discovering new basic knowledge about men and nature" near the bottom of the list.

This conflict between the low priority for funding research and the high confidence that science can solve some major problems of society is interpreted as an indication that the respondents did not understand the questions or did not see the interaction of research and application. The fact that 72% reported positive feelings toward science and technology and ranked scientists as second-most prestigious yet did not favor monetary support for science also shows the public may not see the implications of full appreciation of science. They seem to see science as having great public value but not worthy of public support.

Perhaps such an ambiguous general public opinion about science would be resolved if the public were more literate in the detailed methods of science. Inquiry courses for the general student provide experience in the use of scientific methods. Moreover, to the extent that appreciation of, and confidence in, scientific methods is prerequisite for their retention and later use, the inquiry course could provide an ideal setting in which intelligent public attitudes toward science can develop.

<u>Inquiry</u>

Since there exists a range of views among educators as to what "inquiry" is, a brief review of the types of inquiry must be given, and the nature of the type of inquiry used in the physical science course under study must be described in perspective. One definition of inquiry is given in the document <u>Inquiry Objectives in the Teaching of Biology</u> (McREL, 1969, p.1) as follows:

> "Inquiry...is a set of activities directed towards solving an open number of related problems in which the student has as his principal focus a productive enterprise leading to increased understanding and application."

This definition was written by a joint Mid-continent Regional Educational Laboratory-Biological Sciences Curriculum Study (McREL-BSCS) committee, which undertook a very ambitious inquiry course development program for high school biology, beginning in 1959. Biology was the first science to receive such attention to inquiry methods, but the McREL-BSCS committee, and most later writers regardless of their science backgrounds, attempted to establish an inquiry teaching approach broad enough to cover any science.

A major aspect of the productive enterprise in which the student is involved, in addition to "re-discovering" some of the principles and concepts of science, is that the student is simultaneously directing his own learning, or "learning how to learn" by the inquiry approach.

Beyond this definition, the McREL-BSCS document proposes a rather advanced and formal type of inquiry. On pages 14-19 are outlined the "major factors" in inquiry, and these resemble the traditional pedagogue's view of the "scientific method."

1. Formulating a problem.
2. Formulating hypotheses.
3. Designing a study.
4. Executing the plan of investigation.
5. Interpreting the data or findings.
6. Synthesizing knowledge gained from the investigation.

Although the authors of this outline and some other secondary school curriculum designers propose that students as early as the seventh grade carry out scientific investigations in this sophisticated form, the actual practice in ninth grade and college physical science tends to follow a much less rigorous form of inquiry. For example, in the physical science course that was used in this investigation, the problems to be studied by the student are posed by the laboratory guide, not spontaneously formulated by students, and no laboratory activity requires all the steps of inquiry outlined above.

Recognition of the affective aspects of inquiry courses is also in the McREL-BSCS document (pp.34-37). The following twelve "affective or attitudinal qualities" associated with inquiry behaviors constitute perhaps not a complete list but at least a thorough list. In this list, one example of a behavior that would manifest each attitude is quoted for that attitude:

1. Curiosity--the student expresses a desire to investigate new phenomena.
2. Openness--the student seeks and considers new evidence.
3. Reality orientation--the student does not alter his data.
4. Risk-taking--the student participates in class discussions even if his data differ noticeably from the class mean.
5. Objectivity--the student prefers supported conclusions.
6. Precision--the student seeks definitions of important words.
7. Confidence--the student is willing to take "intuitive leaps."
8. Perseverance--the student pursues a problem until its termination or solution.
9. Satisfaction--the student expresses satisfaction with doing inquiry.
10. Respect for theoretical structures--the student demonstrates awareness of the importance of models, theories, and concepts for relating and organizing new knowledge.
11. Responsibility--the student suggests changes to improve experimental procedure.
12. Consensus and collaboration--the student seeks clarification of another person's point of view.

Development of these qualities in the student in an inquiry course would mean that the student has acquired favorable attitudes toward scientific methods.

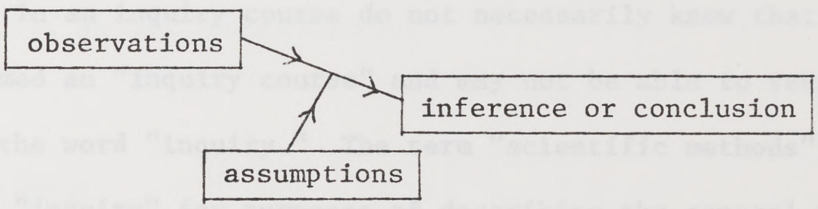
In another view of inquiry, Bruner (1962, pp.81-96) calls it "discovery," which he defines as "rearranging or transforming evidence in such a way that one is enabled to go beyond the evidence so reassembled to new insights." What he has in mind is that inquiry can be taught as an inductive process, and this is emphasized by his claim that inquiry, or discovery, is learned only by attempting to practice it. He proposes that discovery is learned through the "heuristics of discovery," that is, on two levels, by doing and by discovering how it is done. With some exaggeration, Bruner's view has been called the "little scientist" model of learning. Bruner believes that discovery learning can begin in elementary school. In regard to the attitudinal effects of discovery, Bruner "hopes" (1960, p.20) that the "excitement" of discovery will aid favorable attitudes toward working with science subject matter.

Gagne (1963, pp.144-153) takes a much more restricted view of inquiry. He believes that inquiry requires so many prerequisite abilities and facts, most of which could not be learned through inquiry, that only certain advanced high school students could begin to practice inquiry. This view resembles that in the McREL-BSCS statement. Gagne would like to see the learning situation so structured that the inquiring student would propose only valid hypotheses and receive positive reinforcement for them. The student not prepared to inquire would produce mostly "worthless" ideas.

Rutherford (1962, p. 80-84) sees two different views of inquiry as a a teaching method. One point he makes is that inquiry is so much a part of the content of science that inquiry can be taught only if considerable history of science is taught. Curricula such as Harvard Project Physics (HPP) have been designed to take this approach. Rutherford's other point indicates how the views of Bruner and Gagne might be synthesized: inquiry is a valuable teaching technique, but it can be done only with carefully selected laboratory activities, and the form of inquiry in a course will not closely resemble the form of inquiry practiced by scientists. Shulman (1968, p.90) has provided the term "guided discovery" to describe a form of inquiry teaching in which inquiry is practiced in a non-rigorous manner on a planned series of laboratory activities.

The attitudinal effects of an inquiry course could depend strongly on the form of inquiry used in the course, whether it is a "discovery" approach, or a rigorous and structured approach. If a rigorous approach is begun before the students are prepared for it, or if a very elementary "discovery" approach is presented to advanced students, unfavorable attitudinal effects might be found due to inappropriate use of inquiry teaching rather than any basic deficiency in the inquiry approach. Favorable attitudinal effects might be enhanced if the form of inquiry used in a course is matched carefully with the abilities and needs of the students taking the course.

The form of inquiry that actually appears most commonly in college courses would be the form most worthy of study. For example, at the University of Texas, a physical science course enrolling over 900 students in a two-semester sequence is the guided-discovery form that employs a combination of the views of Bruner and Gagne. Its structure can be diagramed thus:



What is to be "observed" by the student is suggested by a laboratory activity manual, although the details of accurate and relevant observation are often left to the student to work out. The activities are in a structured sequence, and usually instructors do not have the students do additional "spontaneous" experiments related to the activities. The students are not asked to form hypotheses preceding the experiments. The outcomes of the activities are intended to suggest some well-known physical concept, or even a law, but the laboratory manual and the instructors usually do not explicitly tell students what they are to "discover." Tangential inferences are often accepted. It is recognized that the students are not prepared for a strict form of inquiry, and so a discovery approach is used. Students are encouraged to freely and intuitively develop inferences from observations made with laboratory equipment, but the laboratory activities done are arranged in a carefully structured sequence rather than being open-ended. Students are not expected to develop their own set of activities. This form of inquiry used might be best termed "guided discovery" (after Shulman). There is a pervasive element of "doing" and manipulation, of "hands-on" experience at all stages of the class work, and a sense of beginning at the beginning with the basic abilities that every college student has. Hereafter, the use of the word "inquiry" in this study will be understood to refer to the "guided discovery" approach. (See Appendix E for the course philosophy as stated to the student.)

It should be clarified in regard to the design of the study, that the students in an inquiry course do not necessarily know that they are in a course termed an "inquiry course" and may not be able to respond to questions using the word "inquiry." The term "scientific methods" can be substituted for "inquiry" for purposes of describing the general processes that occur in an inquiry course. This study is best described as a study of attitudes toward scientific methods rather than a study of attitudes toward inquiry. The students themselves might recognize that they hold attitudes toward scientific methods but be puzzled as to whether they hold attitudes toward inquiry.

Chapter 2. Review of Studies of Attitudinal Effects

Students have attitudes about many things, including science and even scientific methods, before they enter a science course. For the purpose of designing and giving courses, what is of most interest is not the absolute level or amount of student attitude as much as how it changes as a result of the course. Some standard of reference is always needed to establish the amount of the attitude change. Attitude may change during a course and yet the change be completely unrelated to the course. Generally in the studies done so far, not only are pretest and posttest attitude measured in order to ascertain the change, but the change is compared with the change in another similar science course taught in another mode. It is often difficult to locate a course parallel to a physical science inquiry course so that the same population may be assumed to be represented in both courses. This problem can be circumvented if the students who take an additional inquiry course after the first one are compared with those who did not take an additional course.

Changing Attitudes

To expect to measure attitude changes as a result of a student being in an inquiry course, there must be reasons why the nature of the course, or the events in it, change attitudes. It is not enough to say that the student will get an overall impression of the content of the course and it will excite him, leave him untouched, or make him bored, for example.

An attitude change may occur or not occur depending on whether the student is told that an attitude change is expected or what change is expected (Wilkerson, 1972, p.35). Attitude change need not be an intentional goal of a course, yet it will occur in some way, planned or unplanned. If an instructor, or the course structure, explicitly expects certain attitude changes of the student, the instructor cannot be sure those changes will occur. Some students will attempt the change; others may react against the expectation. Methods exist for catalyzing attitude changes covertly, although these are not often used in courses. Since no planned goal of attitude change is pursued in the course under study, nor in the many courses studied by previous investigators to be cited, the issue of planned attitude change need not be considered further.

Secondly, as Wilkerson delineates, student attitudes change when students are in a position of trusting any source (written materials, instructor, etc.) which is suggesting overtly, covertly, or in an unplanned way, the possible attitude changes. In the physical science course studied here, this degree of trust generally holds -- that is, the students are presented with content and/or exercises involving scientific methods and usually recognize they are free to feel whatever they wish about the subject or methods.

A third condition necessary for attitudes to change is that the beliefs that a student holds in regard to a given idea or method not be too "central" in his/her value system. Centrality, in Rokeach's (1968) model of attitudes, refers to how important a given belief is to a person as compared to other beliefs. With the general student, attitudes toward science would not be as central as, for example, attitudes toward close friends. Attitudes toward scientific methods may be even less central and, therefore, more available to change.

Whether planned or not in the course, a principal way that attitudes can change in an inquiry course is through cognitive dissonance (Festinger, 1957). At the level of making observations in the laboratory, the student is continually exposed to conflicting results: other students get unusually different measurements, other students draw conflicting conclusions, a given experiment does not occur the way it "is supposed to," a very unexpected phenomenon is observed. There is a dissonance first at the cognitive level, but then since there is an attitude associated with every piece of cognitive information, there can be affective dissonance. Based on what he has seen--and his personal observation is crucial in inquiry--he wants to believe at least two different things about a situation that had two dissonant results. If a student believes initially that graphing is "worthless," for example, and then discovers that a graph can display a relationship among data that is not apparent in a table of the data, his attitude toward graphing may change favorably.

Cognitive dissonance can also affect more general attitude changes. If a student enters a science course that he does not expect to like, but finds the inquiry course overall to have been rewarding, his general attitude toward science and science courses might change.

Attitude can be changed through positive or negative reinforcement of its behavioral manifestations. A student whose imaginative inferences are derided by class members or always found unacceptable by the instructor soon ceases not only inferring but finding value in inferring. A student whose observations and measurements are <u>always</u> accepted as part of the class' data by also accepting his own estimates of his errors is likely to feel he is capable of making valid observations and will value how the attention to errors supports his efforts.

A recurrent theme in the above description of attitude changes is the instructor. The instructor can provide cognitive dissonance--adversely if the students perceive a gap between what the instructor does and what he believes he is doing as a teacher, and beneficially if the instructor continually exposes students to conflicting situations involving observations of nature which are of such a mild degree the student can cope with resolving them.

Whether he intends to be or not, the instructor is a very important source of reinforcement of behavior. The instructor continually provides small or large rewards for cognitive and affective behavior which he desires and mild punishment for behavior which he does not like. This occurs regardless of whether or not the design of the course includes explicit schedules of reinforcement by the instructor. In an inquiry course, it is desirable that the students gain their own control over what is learned. The student should be able to judge the validity of an hypothesis by what happens when he performs experiments based on it, not be told it is right or wrong or in conflict with the laws of physics by the instructor. The instructor is in a position to easily usurp control over the inquiry situation. Rotter (1969) believes that attitude change comes most readily when the "locus" of control is internal to a person, at least as perceived by that person.

Studies of attitude changes in physical science courses

The search of the literature on previous studies of attitudinal effects of inquiry courses concentrated on college physical science courses. The number of studies involving the inquiry mode was so few that the literature search was broadened to cover attitudinal affects of any physical science course for the general or non-science student. This type of study was found generally to be done as a dissertation study, so a systematic search through all dissertations from 1973 back to 1938 yielded ten studies involving attitude changes in college physical science courses and two in high school physical science. In addition, one study in high school biology was found relevant since it involved both inquiry and affective behaviors. Most of these studies also involved achievement, or knowledge gain, and attention was given to the effects of the courses on achievement, and the interaction of achievement and knowledge.

In regard to the effects of attitude on knowledge or achievement, Krathwohl et al. (1964) reviewed research on the interaction of affective and cognitive objectives and found a variety of results rather than a consistent pattern. For some students, high motivation (positive attitudes) must precede any cognitive learning, whereas for others, decreased attitude follow after cognitive accomplishments. Thus no simple relation between favorable attitudes and high achievement has been established. The studies to be cited do not resolve the issue either.

In reviewing these studies, attention will be paid, in order of importance, to the following factors:

1. Was inquiry used in the course?
2. How was comparison made of attitude changes?
3. Did attitude change more favorably in the "experimental" group (the physical science and/or inquiry course) as compared to a "control" group?
4. Exactly what attitudes were measured--toward science in general, or specifically on inquiry?
5. Were knowledge changes related to attitude?
6. What method of statistical analysis was used?

As will be seen, none of the studies discussed covered all of the criteria holding for the present study--that is, no previous study was on a college physical science course and the course was an inquiry one and attitudes specifically aimed at inquiry or scientific methods were measured.

Studies of physical science courses not involving inquiry were done by Fellers (1972), Grozier (1969), Jones (1969), Wernegreen (1971), and Himaya (1972). Fellers compared attitude changes toward science in a college physical science course for education majors (N=510) with those in an education history course (control, N=180). Measuring attitude with a 55-statement Likert five-position agree-disagree test, he found that the general attitude toward science increased more in the physical science course. Within the physical science course, he compared those who had a laboratory with those who did not and found no significant difference in the attitude changes. No relation of attitude change to sex or grade in the course was found. Analysis was done with group means.

Grozier compared attitude changes in a physical science course for general education students taught by lecture method with and without laboratory. Measuring attitude toward science in general with a Likert-scale instrument, he found that attitude increased over the semester in the course without laboratory and decreased in the course with laboratory. He found no difference in the knowledge gains by the two groups and also found the knowledge gains unrelated to the attitude changes. Analysis of variance was used.

Jones was interested primarily in knowledge changes in a college physical science course without laboratory (N=148). He measured attitude at the beginning of the course only, with a self-developed Reaction

Inventory. In a search for the correlates of knowledge gain, he found initial attitude was correlated positively, and critical thinking was correlated negatively. Again, the attitude of interest was a general one toward science, in accordance with the course's goal of developing scientific literacy. Analysis was by comparison of group means and correlation coefficients. No comparison with another course was made.

Wernegreen compared attitude changes in a physical science course for education students (N=328) with a control group (N=100). Attitude toward "science teacher," "myself," "scientists," "mathematics," "science," and the course were measured, each with a 30-scale semantic differential. He factor analyzed the semantic differential results into 24 supposedly attitudinal dimensions and compared changes in group means on each dimension. On nine dimensions, attitude change was less in the physical science course than in the control group, and on two dimensions, the attitude change was more. Chi-square analyses on background variables yielded no significant relations.

Himaya studied a general education physical science course in an attempt to find what measured variables would predict student changes in understanding, critical thinking, and attitude toward the course. With 26 predictors covering background, personal factors, and instructor, 11.8% of the variance in attitude was predicted using multiple regression, but no significant increase in attitude was found. In predicting other variables, it is interesting to note that the number and type of courses the student had had was usually the best predictor. A significant decrease in critical thinking was also found.

The above group of five studies was limited in several respects but does give an indication of the types of attitude change found in previous research on physical science courses. These courses were taken only by education majors, inquiry was not used in them, and only general measures of attitude were taken. In only one case (Fellers) was attitude unequivocally found to increase, and in one case (Jones), knowledge was found related to attitude. Himaya's use of multiple regression analysis was an interesting approach but, as with the other studies, not done on an inquiry course. Measurement of general attitude toward science in fact seems almost irrelevant to studying an inquiry course, since an inquiry course tends to stress not science in general, which may never even be discussed, but to stress scientific methods as they are performed. However, even in the studies on inquiry-related courses to be discussed below, no attitude measures were taken that focus on scientific methods.

The most common form of college inquiry course in the literature is one based on the curriculum PSNS (Physical Science for Nonscience Students) developed at Rensselaer Polytechnic Institute in 1965 (Wood, 1963 & 1966). PSNS centers around a relatively small amount of subject matter concerned with the physical properties of materials. The textbook consists mainly of simple laboratory exercises which comprise a major way the content and methods of the course are developed, but not the only way. Since lectures and text are also important, the PSNS course generally is "partly" an inquiry course. Three studies of attitude in a PSNS course were found.

Gunsch (1972) compared attitude changes in a PSNS course (N=34) with a physical science course without laboratory (N=60). He measured attitude with the Scientific Attitude Inventory, a Likert-scale test said to cover both "intellectual" and "emotional" attitudes, with alpha=.93. The PSNS course had significantly greater attitude gain, and had greater achievement as measured with his own achievement tests (alpha=.85). The attitude test was not inquiry oriented, however. The relation of attitude and achievement was not studied.

Frangos (1971) measured attitude changes in a PSNS-based course for college freshmen, using four groups (N=23 each), with one group as the experimental group and the other three serving as controls on the tests. Attitude toward science and scientists was measured, as well as knowledge of solid matter and investigatory techniques. Both knowledge and attitude increased, but the two were not related. Analysis of variance on the four groups showed no testing effects.

Diehl (1967) compared attitude changes in a PSNS course with a non-directed laboratory with the changes in a physical science lecture and lab course not based on PSNS. Students were randomly assigned to one of the two courses. Attitude was measured with Rokeach's dogmatism scale, and knowledge was measured with a Facts About Science Test. No significant difference in attitude nor in knowledge was found.

These studies involving PSNS, a partly inquiry curriculum, found a slightly better picture on attitude changes, with two out of three reporting that the attitude changes favored PSNS. However, the attitude measures were not focused on inquiry, but were more general. Studies of PSNS merit more attention to inquiry attitudes.

Three more studies involving physical science courses can be cited:

Oshima (1966) compared a physical science and biology methods course for education students as taught by the lecture-demonstration approach (N=23) with that course as taught by an individual investigation approach using inquiry with the same materials (N=54). Attitude was measured with the Dutton Science Attitude Scale, using 20 agree-disagree scales. Attitude did increase in the individualized approach, and it did not change in the lecture approach, but no significant difference between the two approaches was found. This would seem to indicate that the attitude change in the former approach was also not significant. Achievement was measured with the Read General Science Test and no difference was found between the two approaches. Analysis of variance was used.

Hecht (1970) compared attitude changes in a ninth grade physical science course (N=240) with the changes in a non-science course (N=240). The Introductory Physical Science (IPS) curriculum was used. It is completely an inquiry curriculum (if teachers are trained to use it that way), first developed in 1963. Attitude was measured with six semantic differentials on the concepts "scientists," "science," "research," "science teacher," "science course," and "experiments." Only the latter might have yielded a direct measure of attitude toward inquiry. The overall results were that no increases in attitude were found, as compared with the control group, and on 15 of the 18 dimensions factor analyzed from the semantic differentials, attitude decreased.

Steiner (1970) made the only cited study specifically on attitudes toward inquiry, but this was done in high school biology using the Biological Sciences Curriculum Study (BSCS) curriculum, an inquiry-oriented curriculum. He chose three affective behaviors (curiosity, openness, and responsibility) out of the 18 suggested by McREL and studied the correlations between classroom observations of these behaviors (essentially a non-obtrusive measure) and the teaching practices of the teachers as reported by the students on a checklist. Forty-three classes were studied with a paper measure of affective behaviors, and eight of these were directly observed also. Twenty-seven teacher practices were found significantly related to affective inquiry behaviors. The two practices most directly fostering of inquiry were the teacher's willingness to admit mistakes and telling the students to ask questions behind the statements in the text. No relation was found between affective behaviors and achievement.

Steiner's study differed from all others cited in that no attitude change was measured. Yet because affective behaviors were directly observed, it would be desirable to see this form of study extended to the college level and in physical science.

Two more studies that touch on the use of inquiry can be cited:

Stekel (1970) compared attitude changes in a college physical science course between those students doing experiments from a lab manual and those pursuing open-ended problem selection and design, the latter group being more in the spirit of inquiry. Only attitudes toward science in general were measured, and no difference in attitude change was found.

Wilson (1972) prepared one unit of physical science lab materials for secondary school students with below-average IQ's and other general abilities such as reading. His guidelines for the design of these materials reveal that a rudimentary inquiry-oriented approach was used, although no mention of inquiry was made. The study measured attitude with five semantic differentials on the concepts "science class," "science lab," "science teacher," "the school," and "heat and temperature" (the latter being the topic of the unit). Comparison was made with a control group (also below-average), and the only significant and favorable differences in attitude changes found were attitudes toward science class and science lab. Analysis was by multiple regression, and of the many predictive models formed, the most relevant results to quote are that 24.9% of the variance was accounted for when all pretest attitude measures were used to predict posttest attitude toward science laboratory, and 50.1% of the variance was used in predicting posttest attitude toward heat and temperature. On the one hand, this study is limited because only one unit of work was studied, but on the other hand, this study is of interest because in a science course for the general student, one-half the students must necessarily be "below average."

One criticism of the attitude measures made with the semantic differentials in the above studies is that the results were factored into a large number of dimensions. As Osgood has shown, factor analysis of almost any semantic differential will yield only a few principal factors, and a large number of minor ones. Of the principal factors, only one, the "evaluative," is generally agreed to be a measure of attitude and to be very stable for that purpose. Thus semantic differential results can be properly expressed on only one dimension, "evaluation." This was done in the present study.

Further literature search revealed no other studies encompassing inquiry, attitude, and physical science. If the requirement is made that the attitudinal study focus on inquiry and be done on a college level course, then none of the studies cited covered this requirement. The present study was intended to fill that gap in the literature.

In addition, some improvement in statistical methods is available. All but two of the studies were in terms only of the comparison of group means as tested with t-tests or analysis of variance. The use of multiple regression, as in the studies by Himaya & Milson, not only makes more use of the individual changes experienced by students in courses, but provides a measure of just how much predictive power the variables had. A relation that is to be of practical use must involve a substantial amount of variance. Any sample that is large enough can give statistically significant relations among variables that have no substantial relation. Multiple regression also allows more clearly-stated research questions and more careful analysis of covariance. The use of correlation coefficients cannot be solely relied upon because any correlation coefficient hides a number of different effects. Multiple regression, when the independent and joint contribution of variables to a prediction are separated, can unravel the different effects hidden in a correlation coefficient. Multiple regression consequently was relied upon in a systematic way in the present study.

In the above eight studies involving inquiry courses and general attitude changes toward science, only three studies found favorable attitude changes. It thus was an open question whether the present study would support those three studies that found favorable changes, or the one that found unfavorable change, or the three that found no changes. Since only one study involved attitudes toward inquiry and did not establish change, the present study focused on attitudes toward inquiry. The results of the literature on the relation of attitude and knowledge are just as inconclusive, and this relation was also studied.

Chapter 3. Study Design and Procedure

The population and sample

As described in Chapter 1, this was a study of a sample of the population of those non-science major college students taking a laboratory inquiry course in science. The qualification on the term "inquiry" is understood to be that the course be of the "guided discovery" format, which seems to be the most popular form of inquiry course available for study.

At the University of Texas at Austin, by far the largest number of non-science students taking an inquiry laboratory are those enrolled in physical Science 303 or 304 (hereafter referred to as PS 303 or PS 304), a two-semester sequence. During the time of the study (fall semester 1974), 665 students were in PS 303 and in the next semester, 253 were in PS 304. During the past several years, enrollment in the courses has increased steadily, since they were started in the fall of 1970. In the semester under study, PS 303 had 25 sections, of nominally 24 students each, staffed by 16 instructors. The course meets twice per week for a maximum of two hours per meeting; all meetings are in laboratory rooms.

The subject matter of PS 303, the course which contained the students used in this study, consists of basic physical concepts like mass, time, length, force, and elementary Newtonian mechanics, all presented without mathematics beyond simple arithmetic. The course was originally intended for secondary school science teachers who were going on to teach ninth grade physical science and/or other inquiry science curricula. However, after its inception in 1970, it quickly became a science-requirements elective for students in most non-science fields, but the course was not changed in content or format. Some departments (e.g., architecture) now essentially require it, and the course has acquired the reputation of being an interesting and easy course. More than 75% of the students receive the grade of "A," as would be expected from a course philosophy that emphasizes mastery by all students (see Appendix E).

In the fall of 1974, about 70% of the instructors were graduate teaching assistants, and the remainder were physics faculty members. The instructors vary in their use of inquiry teaching but all follow the manual of experiments (Little, 1973).

For this study, 18 sections, taught by 14 different instructors, were chosen to receive both a pretest and posttest. Four of the remaining sections were given to another investigator for another study, chosen on the basis that he knew the teachers of those sections. One section instructor declined to be in the study, and two sections shifted class hours in a manner that prevented pretesting at the first meeting. There was no reason to expect that the seven sections not in this study were biased or were systematically different in crucial ways from the sections studied.

There were 469 students registered in the 18 sections studied. Of these, 419 (89%) were present at the first day of class and took the pretest. There was no reason to expect that the students absent at that time were biased in a crucial way; they had merely experienced registration delays. Posttests were given during the last two weeks of classes at the instructor's convenience, and 396 students (84%) took the posttest. Those who were absent might have been representative of the absentees throughout the semester, or might have been biased toward disinterest in the course.

It was necessary, for careful analysis of the data gathered, to have pretests and posttests from the same students; in other words, to be able to follow each student's progress from beginning to end of course. On the pretest and posttest forms, student social security numbers were requested and used to match pretest with posttest. When all possible matches were made and all incomplete test materials discarded, exactly 300 students were available as the sample for study (64.0% of total enrollment). Discarding incomplete data could have introduced bias in the form, for example, of discounting those students who could not finish answering the knowledge test. All the sources of bias introduced as the size of the sample was reduced from the total number (665) taking the course to the 300 having complete pretest and posttest data were checked by comparing the distributions and mean scores of the sample of 300 with the distributions and mean scores of the larger numbers of students taking pretest and posttest but not necessarily both.

Comparison of the sample statistics on the variables as presented in Table 3.1 (p.31) and Table 3.2 (p.32) showed an obvious lack of difference between the sample of 300 and the 419 students who took the pretest or the 396 students who took the posttest. Only on the variable sex is the difference noticeable; the proportion of females in the sample of 300 was slightly greater. It was concluded that the sample of 300 studied reflects accurately the larger groups who took the pretest or posttest. Given this close resemblance, it was further inferred that the sample of 300 resembles the total population (469) taking the 18 sections, and resembles the total enrollment (665) in all 25 sections.

However, generalization of the results of the study to the entire population of university undergraduates was contingent upon how well the sample resembled that population, which numbered 33,494 undergraduates, at the University of Texas at Austin in fall 1974. The university distribution

Table 3.1: Distributions on year, major, and sex, compared for all students tested, for the sample studied, and for the university undergraduate population.

Variable	Percent for all pre-tests (N=419)	Percent for matched sample pre- & post-tests (N=300)	Percent for university undergraduates fall 1974 (N=33494)
Year:			
1. Freshman	24%	24%	27%
2. Sophomore	41	41	21
3. Junior	22	23	25
4. Senior	12	12	28
Major:			
Humanities	17	15	5
RTF(Communications)	15	16	8
Social studies	8	7	15
Science & engineering	2	2	26
Nursing & pharmacy	6	7	6
Business	25	23	19
Education	17	19	9
Arts & architecture	4	4	8
General & undetermined	7	7	4
Sex:			
Female	58	62	44.5
Male	42	38	55.5

by year, major, and sex is also given in Table 3.1 and comparison showed some substantial differences. The most obvious difference was that the proportion of science and engineering majors in the sample was 13 times less than the university undergraduate population. This was expected in a course for non-science students. Also, fewer students were majoring in the social studies and the arts in the sample, and more were majoring in the humanities, education, communications, and general studies. Business majors, the largest category in the sample, were approximately representative of the university at large.

Comparison of sample and university distributions by year showed that sophomores are twice as numerous in the sample, and seniors are half as numerous, making the sample somewhat younger than the university undergraduates, as shown by the mean year of 2.22.

Table 3.2: Comparison of variables for all pretests or all posttests and and all matched pre- and posttests.

Variable	All pretests or all posttests	All with matched pre- & posttests
Mean year	2.22 (0.96)	2.22 (0.95)
Mean GPA	2.83 (0.49)	2.89 (0.47)
Mean number science courses required	3.23 (1.5)	3.24 (1.5)
Mean number science courses completed	1.28 (1.4)	1.29 (1.4)
No. students required to take PS304	25 (6%)	13 (4%)
No. students intending to take PS304	192 (48%)*	150 (50%)
Mean score on outside science reading and activity: pretest	4.79 (1.2)	4.72 (1.4)
posttest	4.86 (1.2)*	4.82 (1.5)
Mean score on semantic differential attitude: pretest	61.52 (8.9)	61.69 (8.6)
posttest	59.67 (11.1)*	59.88 (11.1)
Mean score on semantic differential student evaluation of instructor	23.32 (4.4)*	23.24 (4.5)
Mean score on knowledge test: pretest	10.78 (3.4)	10.82 (3.4)
posttest	12.48 (3.6)*	12.94 (3.5)

Notes:
--Standard deviations (sigmas) are given in parentheses for all mean scores.
--Data taken from posttest designated by asterisk (*).
--Percents computed on appropriate pretest count (N=419), posttest count (N=396), or matched count (N=300).

Comparison by sex reinforced the differences between the sample and the university on this variable. Whereas male and female were more nearly equal in number on the whole campus (44.5% female, 55.5% male), the proportion of females taking physical science was substantially larger. One reason was that the proportion of females not in science or engineering (80.7%) was greater than the proportion of males not in science or engineering (68.8%). This sex bias of the sample would be important if sex were established to be an important factor in the results of the study.

The academic ability of the sample as measured by the grade point average could not be compared to the university population because the registrar does not attempt to compute one representative mean for the GPA.

However, the mean of the sample (2.89) represented nearly a "B" average (B=3, C=2). A good guess of the university mean GPA would be slightly lower, perhaps about 2.5.

Basic design of the study

To avoid the problems of finding another comparable population of students with which to compare the students taking the inquiry course, this study was designed to compare the sample internally. Crucial use was made of the fact that some subset of the students in the first inquiry course (PS 303) go on to take the second course (PS 304). Their choice of the second course is a behavioral manifestation of their attitude toward the first course, unless the choice was dictated by outside considerations. Comparison was made between this attitude toward the first inquiry course and the attitudinal changes within that course.

In accord with the goal of the study to ascertain the differences between those students who take the second semester of the inquiry course (PS 304) and those who do not, the design of the study involved assessing the students before and after the first course and finding out who took the second course. In order to take the pretest assessment before any exposure to inquiry has occurred, the pretest was given at the start of the first class meeting. The timing of the posttest is not so crucial and was given anytime in the last two weeks of the semester, long enough before the end to avoid crowding the instructor's regular plans.

The enrollment in PS 304 in spring 1975 was listed on class lists available a few days after classes began. Additions to, and deletions from, the formal enrollment in PS 304 were also periodically obtained. Results on the tracking of the PS 303 students of this study into PS 304 by means of social security numbers are given at the beginning of Chapter 4.

The variables of the study

If one were to try to predict which PS 303 students would continue to PS 304, one would like to know the background of each student, what the students were like before their first exposure to the inquiry course in terms of variables relevant to the expected effects of the course, what experiences the students underwent during the first course, and what departmental requirements might have entered into their decisions. The experiences of a student in a course are usually divided into three domains--the cognitive, the affective, and motor skills. Omitting motor skills, the following list of variables was obtained to cover these areas and have the possibility of being measured. The short name of the variable is given first, followed by a complete description of the variable:

Background variables:

--Year: the year of the student in undergraduate school; also represents the relative age of the student for most of the students in this population.
--Major: the academic area the student reported he/she was registered in at the time of the pretest. Nine areas covering all possible students are used, as listed in Table 3.1 (p.31).
--Sex: male or female.
--GPA: the university grade point average, as reported on the pretest. Not available in the case of freshmen.
--Number of science courses completed: the number completed at the time of the pretest, not counting PS 303 itself. Calculation of courses is obtained by dividing the number of science hours reported by the student by three.
--PS 304 required: whether the student is required by his department to take PS 304 as the second part of six hours in physical science. Possible answers were "yes," "no," and "uncertain."
--PS 304 taken: whether the student took PS 304 or not the next semester. Possible answers were "yes" or "no." Details on the complications in measuring this variable are given in Chapter 4; the basic idea was to use behavior where possible, and when in doubt, to use intent.
--Reason for PS 303: the student was asked to choose one of five listed reasons for taking PS 303 as opposed to other science courses, or write his/her own reason.

Affective variables:

--Pre and post science interest: to obtain this measure, the student was asked how often he/she did outside reading or personal activities involving science. The sum of the two scores is used. It is assumed that science interest is an auxiliary reflection of the student's attitude toward science.

--Pre and post attitude: these measures of attitude were operationally defined as the sums of evaluation scores on a set of semantic differential scales applied to three phrases--"Physical science," "Doing experiments," and "Making inferences from observations," the latter two being aimed at the inquiry aspect of the course.

Cognitive variables:

--Pre and post knowledge: these measures of knowledge are the sums of the numbers of items correct on a 21-item multiple choice test.

Instructor variables:

--Instructor evaluation: this measure came from the sum of four evaluation scales on the semantic differential as applied to the phrase, "Physical science instructor."

--Instructor: this variable merely tells which instructor out of 14 the student had, by assigning one binary variable for each instructor.

The student's grade in the course, an obvious variable, was not felt to be usable for several reasons. The proportion of A's given in the physical science courses is high, 75 to 80%, and approaches 100% in some sections. The variable does not have much spread. Also the basis for giving the grade is very complex and variable, and there is no guarantee it represents either knowledge or attitude or both in a uniform manner. Each instructor decided grades in his/her own way.

Missing data from student responses on the pretest and posttest was accounted for, since otherwise scores of zero are effectively entered as data. In most cases, blanks on the pretest background questions could be filled by referring to the student's posttest and vice versa. In the remaining cases, the mean score was entered for the variables year (one case), GPA (84 cases, mostly freshmen), number of science courses required (34 cases), and number of science courses completed (7 cases). The variable major had a category for blank responses ("undetermined"), and blank responses to the question of taking PS 304 do not arise in the final sample of 300. Blank answers on the knowledge test were scored as "wrong." The five students who did not answer the knowledge test beyond item 20 were not included in the final sample. Blanks on the semantic differentials were very rare and were scored as the midpoint (4) of the scale.

Pretest and posttest design

It was decided to obtain measures of all these variables by asking the students to answer background questions, mark semantic differentials, and take a knowledge test, all on paper. These constituted the pretest. Whatever form was used for the pretest, that identical form was used for the posttest.

The semantic differential is a long established and effective way of measuring levels of attitude (Osgood, et al, 1957), and its indirect approach to attitude is also desired (Wilkerson, 1972). The form of the semantic differential used is in Appendix A. The twelve scales shown are often used and were obtained from Spradlin (1973). For purposes of measuring attitude, only the evaluation scales of the semantic differential are used, and these factor-analyze to be the first, fourth, seventh, and tenth scales (good-bad, unpleasant-pleasant, clean-dirty, worthless-valuable). The other scales serve as dummy scales to distract the student from seeing the semantic differential as a direct measure of attitude. The scales of the semantic differential were to be applied to the phrase printed at the top of the page. Since the goal of this study was to see if an inquiry course affects attitude toward inquiry, inquiry-oriented phrases were used on two of the semantic differentials, and the "physical science" one served a more general purpose in assessing attitude. Four semantic differentials are approximately the tolerance limit of the average student, so the fourth one was reserved for student evaluation of the instructor. That one was placed last so that it would not perturb the student responses on the previous three.

The semantic differentials were conveniently scored by assigning the numbers 1 through 7 to the scale positions, arranging the order so that checking next to the word "good" gives a score of 7 and "bad" scores 1. The "unpleasant-pleasant" and "worthless-valuable" scales must be reversed in scoring. The total attitude score was obtained from the sum of the four evaluation scales on each of three phrases. The maximum score could be 84. If the neutral scale position ("4") is always chosen, the score would be 48. The minimum score could be 12. Similarly, the student evaluation of instructor score is the sum of four scales, with a maximum of 28 and a minimum of 4. Neutral would be 16.

On the pretest, the alpha reliability (Cronbach, 1951; defined in Appendix F) was 0.76 for the semantic differentials used as an attitude measure. On the posttest, the alpha reliability was 0.85. The alpha of the semantic differential used for student evaluation of instructor was 0.85.

A pilot study was done in the summer preceding the main study in order to perfect the form of the questionnaire part of the pretest and to develop a reliable knowledge test. The pilot study resulted in the form of the background questions as shown in Appendix A.

Another question appeared on the pretest in addition to the questions for the variables already listed: the students were asked how many science courses they were required to take. Only the mean was needed for the sample, to be compared with the mean number of science courses completed.

Before all the background data was reduced to punched cards, each student's pretest and posttest background was compared for consistency.

Answers on the pretest were given the most credence where conflicts appeared. In many cases, blank answers could be filled in also during this check. It must be noted that the student's answer to whether he/she would take PS 304 was taken from the posttest. The pretest answer was not used, except to compare with whether the student was required to take PS 304. The student's answer on taking PS 304 was accepted contingent upon his actually being in PS 304.

A search for a physical science inquiry knowledge test that contained questions on inquiry methods as well as on physical science facts revealed the best source to be a 50-item multiple choice test developed by Spradlin and Montague (Spradlin, 1973) for use in such a course for in-service secondary school teachers. They started with a 67-item test and found 50 items that would give an alpha-reliability of 0.95 on their test. (That alpha may be high because many teachers had near-perfect scores on the posttest.) Since this 50-item test covered both PS 303 and PS 304 material, this investigator removed all PS 304 items and any PS 303 items obviously not covered in the course under study. Twenty-five items remained, all having correlations with the total score greater than 0.24 as determined by Spradlin.

This 25-item test was used in a pilot run on two summer PS 303 sections (N=44), with population similar to the fall semester population. The resulting alpha reliability was 0.65, and eight items had correlations with the total less than 0.20. Twelve new items were written and added to the test. This expanded test of 37 items was run on two more summer PS 303 sections (N=43). The alpha reliability was 0.74, and eleven items had correlations with the total less than 0.20 and/or multiple answers.

Correlations less than 0.20 were considered unacceptable and such items were deleted from the test. Three items which were redundant and had low correlations were also deleted.

The final form of the knowledge test had 23 items. When run as a pretest (N=419), the alpha reliability was 0.58, still rather low for a cognitive test, and partly due to the presence of two items with correlations less than zero (items 10 and 11, Appendix A). The low alpha was also due to eight items on which students preferred the wrong answer--to be expected on a knowledge test preceding the course. The two bad items were not used in computing the final knowledge test scores, and the resulting alpha rose to 0.64.

When the 21-item version of the knowledge test was used as a pretest on the sample of 300 students, the alpha was still 0.64. As a posttest, an alpha of 0.68 was obtained, and no item correlation with the total score was less than 0.24. On two items (13 and 17), the students preferred the wrong answer even on the posttest. Complete analysis of the answers to the knowledge test is given in Appendix C and the mean total scores for pretest and posttest are given in Table 3.2.

It should be noted that the answer key was obtained from Spradlin (1973). However, two item answers (41, answer d; 50, answer a) were stated incorrectly in that source. The correct answer was used when one of these items was incorporated in the test used in the present study (as item 15, answer e). Also, where it was clear that an original item had several correct answers, the item either was not used, or the answer choices changed so as to produce one correct answer. (Despite this work, the students still preferred a spread of answers on many posttest items.)

Test scores were obtained merely by summing the number of correct answers. The maximum score could be 21.

In view of the care taken to improve upon the original version of the knowledge test, the persistently low alpha remained inexplicable. Fortunately, the result of the knowledge test was not crucial to the study but rather served as two among many variables.

As shown in Appendix A, the science interest questions consist of two questions in regard to outside science reading and outside science activity to be answered on a frequency scale from "never" to "all the time." These questions are scored by assigning 1 to "never" and so on up to 5 for "all the time." The sum for the two questions is then used. The maximum score could be 10, and the minimum 2. The resulting mean scores for pretest and posttest are given in Table 3.2 (p.32).

Administrative procedure

The cooperation of the instructors was obtained to administer the pre- and posttests. The instructors were given test materials before their classes and they did the actual administration. Appendix B gives the letter to instructors for the pretest. The pilot run showed that no problems should develop with the instructors as administrators, and indeed no problems did. In fact, their cooperation was considerable.

The tests required usually no more than 30 minutes to administer. The instructors administered them in the order of background questionnaire, four semantic differentials, and the knowledge test. The latter made use of optical scan answer sheets.

Statistical methods

Since this study was about the differences between those who take PS 304 and those who do not, the natural, and most important, statistical method was multiple linear regression analysis. In this method, a search was made for those variables which best serve as predictors in a linear equation where "304 taken" is the outcome or criterion. Binary variables like "304 taken" were simply handled by assigning the variable the value 1 for "no" and 2 for "yes." A variable like major, which is not ordered, was handled by a nine-component vector (really, nine variables), each of which represents one major and takes the value "zero" or "one" in accord with the major that is to be designated for that subject.

When all possible variables are used as predictors, the resulting equation is called the "full model." The amount, or weight, of each variable is chosen so as to maximize the total amount of criterion variance explained. Most models can explain only a fraction of the total possible variance. The effect of one or more variables on the prediction can be statistically tested by finding the F-ratio of the amount of variance predicted by the full model as compared with the amount when certain predictors are not included in the model. When using multiple component vectors to represent variables such as major or instructor, care was taken that one degree of freedom was dropped for each multiple representation. For example, a nine-component vector for major represents eight independent degrees of freedom, because once it is known a student was not in eight of the majors, then he must have been in the ninth.

In predicting the decision to take PS 304, the search for the relevant variables started from a position of ignorance. The best statistical method is to use commonality analysis (Veldman, 1975; Kerlinger & Pedhazur, 1973), where the variables are taken as sets, e.g., background, attitude, knowledge, and instructor. Models are formed using sets of variables as the basic predictors. Once it is found which sets are important, the sets can be broken into their component variables to isolate which variables are important predictors. Commonality analysis separates the independent from the joint effects of the predictors, so as to find out what groupings of variables have predictive value in common. No statistical tests of significance are available for the joint effects, but their importance can usually be inferred from the amount of variance being accounted for, a result provided by the analysis.

Once the important predictors of taking PS 304 were found, their relations to any possible changes in knowledge, attitude, and science interest during PS 303 were explored. The most confounding effects checked in the analysis were the effect of instructor, and the effect of being required to take PS 304.

Chapter 4. Results

The decision to take the second course

On their posttest questionnaire, the students answered the question of whether they would take PS 304 with "yes," "no," "maybe," or "blank." The guideline for reducing these answers to a clear "yes" or "no" was to consider the student's intent, since it is an attitudinal effect that was sought.

For the 300 students for whom pretest and posttest data were available, 150 (50%) said they would take PS 304. Of those, 118 (39%) actually took PS 304 the next semester and 32 (11%) did not. This proportion (39%) reflected very closely the ratio of the total in PS 303 in the fall (665 students) to the total taking PS 304 in the next semester (spring; 253 students, or 38%), thus confirming again the representativeness of the sample Considering the fact that on the pretest 35% said they would take PS 304 whereas by posttest 50% had decided to, and considering the fact that saying so would seem to be sufficient intent for the purposes of this study, and considering the fact that some students take PS 304 more than one semester later, this study was justified in assuming that all who said they would take PS 304 would eventually do so. In going through the PS 304 enrollment, it was found that 43% of the PS 304 students had not been in the 18 sections of PS 303 studied the previous semester. The conclusion is supported that some PS 304 students come from previous years of PS 303, an estimated 30% of them.

Of the 25 students who were uncertain about taking PS 304, 24 did not take PS 304; these students thus had given an effective "no" answer and were so considered from this point on.

Of the 150 students (50%) who said they would not take PS 304, only 4 (1.3%) did indeed take PS 304, so that the "no" answers can be taken as accurate.

Of the 300 students, only 13 (4%) said they were required to take PS 304 and 8 (2%) were uncertain. Clearly, being required to take PS 304 could not be an important predictor of who takes PS 304. Reading the various answers on the questionnaires showed confusion among the students as to what number of science hours were required, whether PS 304 was required, and how many hours of science they had completed. This confusion was inferred from inconsistencies among the answers to the questions of how much science was required, why PS 303 was taken, was PS 304 required, and how much science had been taken. For example, an occasional student would say that PS 304 was required but that he was not going to take it.

Research questions

In this study, the behavior of students after an inquiry course (PS 303) as manifested in their decisions to take or not take PS 304 (the second semester of the same type of course) was proposed as an unobtrusive measure of the overall attitudinal effect of the inquiry course. The research questions were aimed primarily at eliciting what measurable factors in regard to student background and student experience in PS 303 might determine their decision to take PS 304. Secondly, the structure of their experience in PS 303 was explored in an attempt to clarify how that experience did or did not affect their decision.

The variables of the study fell into five sets:

Set 1 (2 variables): the pretest and posttest attitude as measured by the four evaluation scales of the semantic differential on each of three concepts related to inquiry learning.

Set 2 (2 variables): the pretest and posttest knowledge of physical science as measured on a 21-item multiple choice test.

Set 3 (2 variables): the students' reported outside interests in science as measured on pretest and posttest by questions on how much science reading and science activities were done outside courses.

Set 4 (14 variables): student background--year, major (represented by 9 binary vectors), sex, GPA, number of science courses completed, and whether a student's department actually requires that he take PS 304.

Set 5 (15 variables): student evaluation of instructor as measured on posttest by four evaluation scales of the semantic differential, and 14 binary vectors telling which of 14 instructors the student had.

The main questions of the study were:

1. Which of the above five sets of variables, or combinations thereof, significantly predict for each student whether he or she takes PS 304?
2. Within set 1, does change in attitude during the semester, or the level of attitude itself, predict the student's decision to take PS 304?
3. Within set 2, does change in knowledge during the semester, or level of knowledge itself, predict the student's decision?
4. Within set 3, does change in science interest during the semester, or the level of interest, predict the decision?
5. Within set 4, which background variables are significant predictors of the decision?
6. Within set 5, do any particular instructors affect the decision, does any particular instructor receive an unusually different evaluation, and does instructor evaluation predict the decision?

Following the results of the above questions, certain variables or sets of variables were found to be important. Followup questions as to their interrelations were asked, particularly as to how attitude and/or knowledge levels or changes in the PS 303 course were affected. In particular:

7. Did attitude change during PS 303?
8. Did knowledge change during PS 303?
9. Did science interest change during PS 303?
10. What variables were related to attitude change?
11. What variables were related to knowledge change?
12. What variables were related to instructor evaluation by students?

Commonality analysis

Each of the following research questions was answered by casting the question in the form of a null hypothesis and subjecting the hypothesis to a standard statistical test.

Answering the first question of which sets of variables significantly predict student decision to take PS 304 was done with this null hypothesis:

> In predicting the decision to take PS 304, there is no significant difference between the amount of variance accounted for when all variables are used as predictors (full model) and the amount of variance accounted for in each of five models formed when the five various sets of variables are each dropped in turn from the full model.

Commonality analysis, which distinguishes between the independent and the joint (or common) contributions to the variance by the sets, gave the results shown in Table 4.1 for the five versions of the null hypothesis, as each set of variables was dropped from the model. An F-test was performed in each case and the resulting probability that the F-ratio obtained was due to chance is shown. How the sets of variables overlap in predicting the criterion is diagramed in Figure 4.1 (p.47).

Table 4.1: Commonality analysis on all sets of variables: independent contributions.

Variable set	Percent of predictable variance	Percent of total variance	Probability
Attitude	1.41%	0.34%	0.5578
Knowledge	6.86	1.66	0.0556
Science interest	0.68	0.16	0.7568
Background	51.50	12.46	0.0003*
Instructor	25.49	6.17	0.1320*
Totals	85.94%	20.79%	

Note: The asterisk (*) indicates that the probability is a slight overestimate (more conservative than necessary). The independent degrees of freedom for the numerator of the F-ratio was one less than the nine components of the vector for major and the 14 components for instructor.

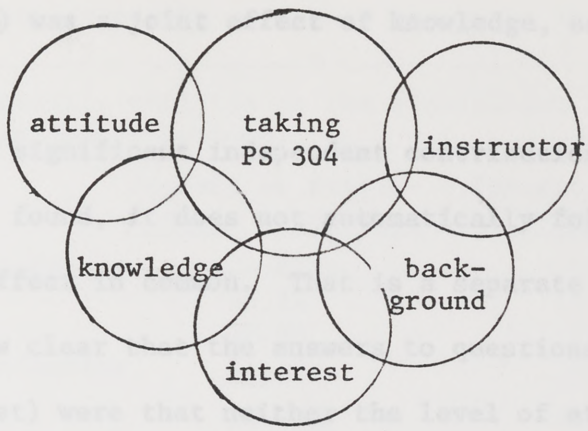


Figure 4.1: Independent and joint contributions by sets of variables.

Note: Joint contributions are shown as overlaps. (Not all overlaps can be shown in two dimensions.)

Wherever the probability for a set was less than 0.05, the null hypothesis was not rejected, and an alternative conclusion was drawn. Otherwise, it was rejected. The total variance (R-squared) predicted by the full model was 0.2420, or 24.20%. The independent contributions accounted for 20.79%, or 85.94% of all the variance that can be predicted by the model. The joint contributions must account for the remaining 14.06%. The probabilities listed are for F-tests performed on the percent of total variance as compared to the total of 24.20%.

Only the knowledge and background sets gave significant contributions, although instructor effects received some further attention.

The above sets of variables can be combined in 26 ways to produce joint effects of two, three, four, or all five sets of variables. Only one combination, knowledge and background, made a joint contribution of greater than 1% of the total variance, which, as will be seen later, is a reasonable criterion for considering contributions to a model prediction.

That pair contributed 1.71% of the total variance. The next largest contributor (0.64%) was a joint effect of knowledge, science interest, and background.

Although significant independent contributions of knowledge and background were found, it does not automatically follow that they will have a significant effect in common. That is a separate finding.

It is now clear that the answers to questions 2 (attitude) and 4 (science interest) were that neither the level of attitude or interest nor the changes in attitude or interest are significant predictors. It follows from the fact that, if pretest and posttest attitude together were not significant, then no linear combination of them, including their difference, can be significant.

Questions 3 (knowledge), 5 (background), and 6 (instructor) were pursued further by commonality analysis on the following model (Fig. 4.2), where the knowledge and instructor sets have been subdivided, dropping pretest knowledge to test the effect of posttest knowledge alone.

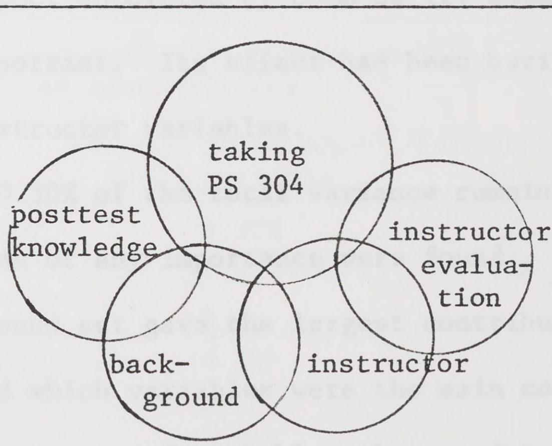


Figure 4.2: Commonality analysis: non-significant sets of variables dropped.

When the attitude and interest variables and pretest knowledge were dropped from the model, the total amount of variance of the criterion (taking PS 304) explained by the model was reduced to 22.99%. The independent contributions by the variables are given in Table 4.2.

Table 4.2: Commonality analysis on the significant sets of variables: independent contributions.

Variable set	Percent of predictable variance	Percent of total variance	Probability
Posttest knowledge	4.67%	1.07%	0.0509
Background	60.21	13.84	0.0001
Instructor	22.55	5.18	0.2105
Instructor evaluation	11.29	2.60	0.0032
Totals	98.72%	22.69%	

These results established that posttest knowledge was a barely significant predictor, with about the same predictive effect as pretest and posttest knowledge together. Therefore, posttest knowledge had little predictive value independent of pretest knowledge, and it was inferred that the level of knowledge, not a change in knowledge, was the significant predictor. In further models, pretest knowledge was restored to the set of knowledge variables, and the set interpreted as measuring the level of knowledge.

Another aspect uncovered by commonality analysis was that instructor evaluation was important. Its effect had been buried in the 14 degrees of freedom of the instructor variables.

With only 0.30% of the total variance remaining out of 22.99%, no joint contributions of any importance were found.

The background set gave the largest contribution, and it was divided into parts to find which variables were the main contributors. Figure 4.3 (p.50) shows the new sets of variables that go into the full model predicting taking PS 304.

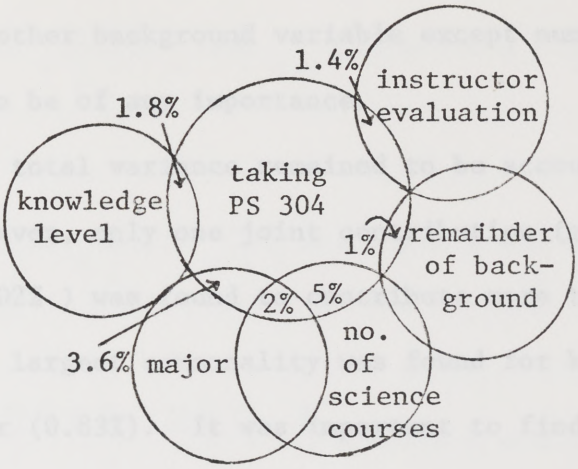


Figure 4.3: Subdividing the background set and including all significant contributors.

Note: Portions of variance independently and jointly explained are represented here.

When all irrelevant variables were dropped from the model, commonality analysis gave the independent contributions for the remaining variables shown in Table 4.3. The amount of total variance of the criterion explained was 18.73%.

Table 4.3: Commonality analysis on the remaining sets of variables: independent contributions.

Variable set	Percent of predictable variance	Percent of total variance	Probability
Level of knowledge	9.72%	1.82%	0.0426
Major	19.43	3.64	0.1855*
Number of science courses completed	26.29	4.92	0.0002
Remainder of background	5.31	0.99	0.4890
Instructor evaluation	7.51	1.41	0.0262
Totals	68.26%	12.78%	

Note: The asterisk (*) denotes a slight overestimate, as explained in Table 4.1 (p.46).

Once major, which carried eight degrees of freedom, was removed from the background set, no other background variable except number of science courses was found to be of any importance.

5.95% of the total variance remained to be accounted for, in joint contributions. However, only one joint contribution (major and number of science courses 2.02%) was found to contribute more than 1% of the total variance. The next largest commonality was found for knowledge, major, and background remainder (0.83%). It was important to find that number of science courses had little in common with instructor evaluation (0.43%).

To summarize the results so far, the independent and joint contributions of the variables predicting the taking of PS 304 were approximately as depicted in Figure 4.3 (p. 50). Only the level of knowledge, number of science courses completed, and instructor evaluation were important predictors. Change in knowledge, attitude level or change, major, science interest, and which instructor all were not significant predictors.

It must be emphasized that the amount of variance predicted in the full model (24.2%) was not very large. In other words, there were no variables, or combinations of variables, that accounted for more than about one-fourth of the student decisions to take the next course. In this perspective, only the number of science courses completed can be said to be a substantial predictor of the decision, and it was not an "important" variable in the same sense attitude or knowledge would be.

Correlation coefficients were computed between all pairs of variables in the study. Selected and significant ones of interest are displayed in Table 4.4 (p.52). It must be noted that correlation coefficients contain both independent and joint relationships between variables, so that they must be interpreted with some caution.

Table 4.4: Correlation coefficients between selected pairs of variables.

Variable \ Variable:	GPA	No. science courses	PS304 taken	Pretest attitude	Posttest attitude	Pretest knowledge	Posttest knowledge	Eval. of instructor
Year	x	.59	-.10	x	x	x	x	x
Major: humanities	.22	.14	x	x	x	x	x	x
communications	x	x	x	x	x	x	x	x
social studies	x	.13	x	x	.10	x	x	x
business	x	-.10	-.13	x	x	.12	.12	x
education	x	-.30	.24	x	x	-.23	-.16	.11
GPA	1	x	x	x	x	.18	.17	x
No. science courses	x	1	-.31	x	x	.16	x	-.13
PS 304 taken	x	-.31	1	x	x	-.20	-.14	.16
Pre science interest	x	.16	x	.28	.21	.21	.13	.11
Post science interest	x	.17	x	.24	.29	.25	.19	x
Pre attitude	x	x	x	1	.42	.18	x	.21
Post attitude	x	x	x	.41	1	.13	.12	.53
Pre knowledge	.18	.16	-.20	.18	.13	1	.59	x
Post knowledge	.17		-.14	x	.12	.59	1	x
Eval. of instructor	x	-.13	.16	.20	.53	x	x	1
Instructor: 1	x	x	x	x	x	x	x	.15
2	-.11	.11	x	x	x	x	x	x
3	x	x	x	x	x	x	x	x
4	x	-.13	x	x	x	x	x	x
5	x	x	x	x	x	x	x	x
6	x	x	x	-.17	-.19	x	-.14	x
7	x	x	x	x	x	x	x	.12
8	x	.16	x	x	x	x	x	-.13
9	x	x	x	x	-.14	x	x	-.33
10	x	.17	x	x	x	x	x	-.12
11	x	x	x	x	x	x	.21	x
12	x	-.16	x	x	x	x	x	x
13	x	x	x	.10	x	x	x	x
14	.10	x	x	x	.16	x	x	.19

Note: Only significant correlations are shown. Correlation r must be greater than 0.097 to be significant at 0.05 level, by a t-test for r for N=300.

In regard to the number of science courses, the correlation between that variable and the decision to take PS 304 was -0.31, quite significant and negative. Those who took PS 304 had completed the fewest number of science courses. Further information is available if it is recalled from Table 3.2 (p.32) that a mean of 3.24 science courses was required of the students, and a mean of 1.29 science courses had been completed before PS 303. The conclusion was clear that the average student needed two more courses, one being PS 303 and the other logically being PS 304 unless the student had reason for deciding against PS 304.

In regard to major, Table 4.4 (p.52) indicates that business students were least likely to take PS 304 after PS 303 and education students were most likely. (Most of the education majors were in elementary education or special education. All secondary education students take a physical science course offered in their department.)

When correlations between the instructors and the students' evaluations of them were examined, it was found that Instructor 9 received a substantially lower evaluation than the others. However, this instructor was not an anomalous contributor of variance to the prediction of taking PS 304, as that instructor was not significantly related to a decision to take PS 304 (or not take it, as a large negative correlation would show). As a further check, data from this instructor's section was temporarily dropped from the model, but no significant changes in the effects of any predictors were found. In particular, student evaluation of instructor remained the same significant predictor of the decision.

Changes during the inquiry course

Aside from science course requirements and possibly the student evaluation of instructor, it has not yet been explained why 50% of the

PS 303 students continued on to take PS 304. Only 4% had reported that they were required to take PS 304, so the rest must have chosen it voluntarily from among the other possible science courses. Their experiences during PS 303 were examined in order to provide further clarification as to why they might have chosen PS 304.

Returning to the research questions, the following null hypotheses were formed to test the significance of the students' various experiences during PS 303:

(7) There is no significant difference between the mean of the scores on the pretest semantic differential measure of attitude (12 evaluation scales) and the mean on the posttest).

A t-test for correlated distributions (Guilford, 1973) was performed, giving t=-2.90, significant beyond the 0.01 level. The implication was that attitude changed in the negative or unfavorable direction during the semester.

(8) There is no significant difference between the mean of the scores on the pretest of knowledge and the mean on the posttest.

Again a t-test for correlated distributions gave t=11.80, significant beyond the 0.01 level. The conclusion was that knowledge increased, in fact, about two test items out of 21 items on a multiple choice test (pretest mean=10.82, posttest mean=12.94, perfect score=21).

(9) There is no significant difference between the mean of the scores on outside reading and science activity at the beginning of the semester and the mean at the end.

A t-test on the resulting t=1.49 was not significant at the 0.05 level. This null hypothesis was not rejected and the conclusion was that no change in interest has been shown.

At this point, it was clear that some changes in attitude and knowledge occurred during the PS 303 course, yet had no demonstrable effect on taking the next course. It was helpful to check the relation between the students' evaluations of their instructors and their knowledge and/or attitude and changes thereof. A model was formed using all the variables to predict posttest attitude and the following null hypothesis tested:

(10) In predicting posttest attitude, there is no significant difference between the amount of variance accounted for in the full model and the amount in the reduced model with instructor evaluation dropped.

The resulting F-ratio of the variances was 80.56, significant beyond the 0.0001 level. The null hypothesis was rejected and the conclusion was that those students who rated inquiry high on the semantic differential also rated their instructor high.

In the full model, using all variables to predict the posttest attitude of students, the most important predictor was found to be instructor evaluation, accounting for 28.5% of the variance (independent and joint) out of the total of 50.3% that this model can account for. (This model, with 50.3% of total variance, was rather substantial as a predictive model.)

Examination of the correlations of other variables with posttest attitude in Table 4.4 (p.52) showed that the variables major, science interest, pretest attitude, knowledge, and instructor merited further attention. When each of these variables was successively dropped from the full model and an F-test made of the variance of the reduced model as compared with the full, the results shown in Table 4.5 (p.56) were obtained.

Major was a significant predictor of posttest attitude, with social studies majors showing the most change from pretest to posttest, as examination of the correlations between major and attitude in Table 4.4 (p. 52) shows.

Table 4.5: Predicting posttest attitude.

Variable tested	Percent of total variance dropped	F-ratio	Probability
Major	3.3	2.209	0.0269
Pretest science int.	.16	.816	0.3703
Posttest science int.	1.9	4.877	0.0085
Pretest attitude	7.3	37.449	0.0001
Pretest knowledge	.03	.133	0.7165
Posttest knowledge	.04	.209	0.6526
Instructor	4.2	1.661	0.0688

Note: In these models, degrees of freedom were adjusted to represent the correct number of independent degrees of freedom: eight for major and 13 for instructor.

Posttest science interest was an unexpectedly significant predictor of posttest attitude, or at least it was clear that those students with higher attitudes also reported more interest in science at the end of the course.

Pretest attitude was, of course, a significant predictor of posttest attitude, but not the main predictor, since attitude did significantly change. (It should be noted that all these models automatically account for the covariance of variables not explicitly dropped from models.)

Knowledge was not a significant predictor of posttest attitude and therefore of attitude change.

Despite some significant individual correlations, the instructors as a whole were not a clearly significant predictor of posttest attitude, although they were on the borderline. Examination of correlations between instructors and attitude showed four instructors that stand out somewhat from the rest. Instructor 6 had a class whose overall attitude started low on the pretest and stayed relatively low on the posttest. Instructors 9 and 13 had classes whose overall attitude was more unfavorable on the posttest, and Instructor 14's class had an overall favorable shift. These fluctuations could be expected by chance among a group of 14 instructors.

It may be seen at this point a hint as to why attitude had no demonstrable effect on the decision to take PS 304. Attitude overall was decreasing, yet those with higher posttest attitudes rated instructor higher.

The relation of instructor evaluation and knowledge was elicited with this null hypothesis:

(11, 12) In predicting posttest knowledge, there is no significant difference between the amount of variance accounted for in the full model and the amount when instructor evaluation is dropped from the model.

The F-ratio of the variances was 1.387 ($p=0.2381$, not significant). This null hypothesis was not rejected, and it was concluded that the gain in knowledge was not related to evaluation of the instructor. In fact, the overall level of knowledge also was not related, since pretest knowledge was the most potent predictor of posttest knowledge (34% out of 43%).

Examination of the correlations between knowledge and other variables in Table 4.4 (p.52) showed that major, GPA, taking PS 304, science interest, posttest attitude, or instructor might be related to a prediction of student posttest knowledge. However, forming a predictive model of knowledge (which accounted for 43.4% of the total variance) and making F-tests with models formed by dropping these variables, showed only instructor to be significant. Instructor accounted for 5.5% of the variance and yielded an F-ratio of 2.024, significant at 0.0190. Thus, although some other variables had significant correlations with knowledge, they were not strong enough to be significantly predictive after all covariate effects were accounted for.

The table of correlations shows that one major (education) consistently received the lowest scores on both the pretest and posttest of knowledge. Only two instructors, 6 and 11, stood out as having classes with noticeable changes in knowledge. One (11) was not the same instructor as the ones mentioned in regard to attitude or evaluation, and the effect may be only spurious.

Reason for taking PS 303

Data on one other variable was obtained: the student's report of his/her reason for taking PS 303, selected from among reasons such as "recommended by a friend," "required," "easy course," "no lectures," "sounded interesting," or written in a blank. Occasional written answers were "felt needed the course," "no homework," "advised to take it," or "needed a course." Because the students found much overlap among these various reasons, or specified more than one, the reasons were not sufficiently quantifiable to form a multiple vector for use in the models. However, overlap among the reasons does not prohibit the use of chi-square analysis to see if there is any relation between the reason and the decision to take PS 304.

The frequencies with which given reasons were chosen by those who did and did not take PS 304 are shown in Table 4.6 (p.59). Chi-square analysis was done for this table of frequencies, and the resulting probability was 0.0389 that the difference in reason between those taking PS 304 and those not taking it was not significant. The conclusion was that those who do take PS 304 differed from those who did not as to their reasons for taking the first inquiry course, PS 303. Although "recommended

Table 4.6: Reasons for taking PS 303.

Stated reason	Frequency for students taking PS 304	Frequency for students not taking PS 304
"Recommended by a friend"	76	79
"Required"	21	17
"Easy course"	11	24
"Sounded interesting"	30	15
"Advised to take it"	6	5
Remainder	4	8

by a friend" was the most frequent reason, it did not seem to differentiate those who took PS 304 from those who did not. The "easy" and "interesting" reasons, although less frequent, seemed to discriminate better between those who took PS 304 and those who did not. The results of examining their reasons seemed to support the idea that students chose PS 304 because they "liked" PS 303, perhaps partially for its novel format.

Chapter 5. Conclusions and Recommendations

Before conclusions are inferred from the results of Chapter 4, it should be pointed out that the analysis has been done on data in the form of quasi-normal distributions of measures taken on a sample of people considered as a whole in various ways. The questions studied have been of the form, "Does the sample as a whole change in a certain way?" or "Does the sample as a whole have a relation between certain variables?" The results are in the form of slight average shifts that are or are not found with certain variables.

By the nature of statistical distributions, there will be some students in the sample for whom the opposite of the general result occurred, or for whom something happened when nothing happened for the sample as a whole. Real effects of a course occur with individuals regardless of whether or not real effects occurred with the sample as a whole. In studies there is a choice of whether individuals or whole samples will be studied. The approach to be used depends on the goal of the educational study.

In this study the primary goal has been to demonstrate some overall effect of the inquiry teaching method on attitude. Focusing on the individuals to find out what the effects on each person were without regard to the effects on others would be a different approach and for some purposes more desirable. To some extent the present use of multiple linear regression and the tracking of each individual's changes during the course were stages toward making fuller use of the effects of a course on individuals even though the results were summed over all students. Multiple linear regression also provides a measure (R-squared) of how much individuals contributed to summed effects as compared to individual fluctuations.

Conclusions

The general conclusion was that no attitudinal effect of the inquiry course was demonstrated. The attitude of the students as manifested in their choice of taking or not taking a second inquiry course following the first was not found to be related to their attitude level or attitude changes in the first course. The only variables of the study that were found significantly related to the student decision to take a second inquiry course were a) the number of science courses the student had completed, b) the evaluation the student made of the instructor, and c) the level of knowledge the student had (but not the gain in knowledge). These are in descending order of importance to the decision.

Number of science courses was found inversely related to the decision (correlation was -0.31). The more science courses the student had had, the less likely he was to take the second inquiry course. Student evaluation of the instructor was directly related to the decision; that is, the more favorably the student viewed the instructor, no matter which instructor he/she was, the more likely the student would take the second inquiry course. The level of knowledge was also positively related to the decision in that students who knew more were more likely to take the second course.

In regard to attitude as measured with the semantic differentials within the inquiry course, the mean change in attitude was found to be significantly negative or unfavorable. The exact amount of the change is meaningful only in terms of the exact measurement procedures as used in this study. From Table 3.2 (p. 32) the change was -1.81 relative to a pretest score of 61.69, not a large change.

The attitude change during the inquiry course was found to be related to several other variables. The most significant relation was with student evaluation of instructor, again independently of who the particular instructors were. That relation accounted for over half (28.5%) of the total variance (50.3%) explained in posttest attitude. Other variables found significantly related to attitude change were major, posttest science interest, and possibly who the instructor was. The effect of the instructor bordered on insignificance.

In comparing the attitudinal effects found in this study with those found in previous studies of physical science courses, care must be taken to discriminate between previous courses that used inquiry and ones that did not. As discussed in Chapter 2, two out of five studies found attitude to decrease in non-inquiry physical science courses, as did this study. Two other studies found no change, and one found an increase in attitude. The point of commonality between this study and those five previous studies is that all involved physical science courses with some amount of laboratory activity. The earlier results of either a negative change or no change seem to resemble the small negative change found in this study. This comparison is strongly qualified by the fact that earlier studies measured only general attitude toward science, not attitude toward inquiry, and the fact that the populations of the studies consisted of education students, not general college students.

When compared with previous studies of inquiry courses, the attitudinal results of the present study seemed to be in accord with the one study out of eight that found unfavorable changes in attitude. However, the comparison was made very cautiously because only the present study used attitude measures explicitly on some aspects of inquiry.

The fact that 50% of the students in the inquiry course voluntarily wanted to take a second inquiry course constitutes some attitudinal affirmation of the course. That this student behavior was not related to the decreasing attitude toward scientific methods within the first course was an indication that other considerations besides the inquiry aspect were determinants of student attitude toward the course. This general attitudinal affirmation of the inquiry course was supported by three previous studies that found favorable general attitude changes with inquiry courses.

Physical science knowledge, as measured with a 21-item multiple choice test, increased from a mean of 10.8 items correct on the pretest to a mean of 12.9 items correct on the posttest, a gain of 2.1 items out of 21, or 10%, not very large but still significant beyond the 0.01 level. For comparison, if the knowledge test had been answered only by guessing, the result would have been a binomial distribution of scores with a mean of about four items correct and a sigma of 1.8 . The variance actually obtained (sigma=3.4 on pretest, 3.5 on posttest) is larger and indicates there was a very wide range of cognitive ability on the physical science test, and that range of ability was unchanged during the course. Since the mean scores on the test were approximately half of the total possible score, the test had nearly an ideal level of difficulty for purposes of discriminating among the students as to their ability with physical science knowledge. However, to the extent to which the test reflected true course content and could be mastered by all students on the posttest, the posttest mean fell far short of the level of full mastery.

The only variable found related to knowledge gain was the instructor. The knowledge gain was greater if a student had certain instructors, and less or negative under other instructors.

The only way in which knowledge level or gain was found related to attitude was in the finding that students with a greater level of knowledge (but not gain) were more likely to demonstrate favorable attitude toward the course by taking a second such course. Otherwise, within the first course, attitude level or changes were not found related to knowledge level or gain. The tendency in previous studies has also been to find no relation of knowledge to attitude. The relation of knowledge level to decision to take a second inquiry course was independent of the effects of other variables such as GPA, major, number of science courses, or other variables that might influence the decision. However, the relation was not strong enough to warrant being called an important finding. A possible interpretation of the relation was that those students who perceived that they were less knowledgeable than other students during the first course were less inclined to take another of the same type of course and risk learning less and/or a lower grade. Conversely, those students who already knew more during the first course may have retained confidence to take the next one.

Aside from these considerations, the main reason that was found for students taking a second inquiry course was that they seemed to need it to complete their science course requirements.

There was some indication of an instructor effect. This was best seen by looking at the results of student evaluation of instructor, since the effect of which instructor a student had was spread among 13 independent degrees of freedom and would not show significance unless particular instructors had a very strong effect. Student evaluation of instructor was found significantly related to the decision to take a second inquiry course, and this was an independent effect of that variable. Student evaluation of instructor was also found to be the most important predictor of posttest attitude. Since there was no significant relation between posttest attitude and the decision, there seems to have been an instructor effect in the two different attitudinal areas--in the attitude toward the first inquiry course, and in the attitude toward scientific methods.

Major was found related to attitude change in the first course, with social studies students having the largest favorable change in attitude. The effect was not large because the proportion (7%) of such students in the course was so small, but the effect is expected since social studies students might be more inquiry-oriented by virtue of their field than any other major taking the course. No other particular major was found related to attitude change. In regard to the decision to take the second course, elementary education students were most likely to take the second course, although they were not required to. Perhaps for their field they perceived a need for additional science experience, especially with the inquiry method that is now being used in some elementary science curricula. Business students were least likely to take a second course.

The remaining results of the study were important only in the sense of being null results. No relation was found between the decision to take a second inquiry course and background variables such as year, sex, or GPA. These variables also were not found significantly related to attitude changes during the first course, nor to knowledge gain, nor to the level of attitude or knowledge.

The results of this study were obtained in a particular lower division college physical science inquiry course at a particular university, with a particular sample of students. Generalizing the results was contingent on consideration of the characteristics of the population represented by the sample and on the nature of the inquiry course studied. The sample was shown to be well-representative of the total enrollment in the course. The resemblance of the total enrollment to the university undergraduate population was not as strong. When allowance for the absence of science and engineering majors in a course for non-science students is made, the university undergraduate population was found to differ from the sample, with more of some majors and less of others. Year was biased toward the older side as compared with the university undergraduate freshmen and sophomores. Sex was biased in favor of females, and GPA was biased toward being slightly higher in the sample. The population choosing to take the inquiry course was clearly selective in some ways. However, it was reasonable to suppose that similar selection would occur at other universities with similarly diverse populations. Then the results of this study would be applicable to similar physical science inquiry courses at other universities and colleges.

More caution would be needed in generalizing to community college populations, because on the one hand, the community college student is more oriented toward practical application than fundamental ideas such as inquiry or physical science concepts. But on the other hand, the community college student tends to feel disadvantaged in traditional classroom situations and might respond better to the non-abstract approach of inquiry and the personal attention from instructors. Supportive of this, instructor effect was found to be important.

Generalizing the results to the secondary school level seemed unwarranted because of considerable differences in age and experience, and because the university population is a non-random selection of the high school population. Less than half of secondary school students continue to college, and secondary school classes include that part of the population which consists of potential science and engineering majors, a group not included in the present study.

Generalizing the results of this study to inquiry courses in other science subject matters at colleges and universities also warrants considerable caution. There was no evidence that the particular physical science content was unimportant both in student decisions to take more inquiry courses and in changes in student attitudes toward using scientific methods on that content.

In a limited way the results could be generalized to college physical science courses not using inquiry but having a laboratory-oriented approach. Since attitude toward inquiry was not found related to the taking of a second course in the sequence, some non-inquiry aspects of the first course might predict the taking of a second course. The principal effect might be that students who need one more science course and who liked the first laboratory course elect to take the second. Attitude toward the instructor may also predict the taking of a second laboratory course. Since that attitudinal effect was independent of particular instructors, it might be that the personal contact with instructors in the first course influenced some to take another course with that type of interaction with instructors.

If it is assumed, with some justification, that the general student would often prefer to have no science requirements, then the finding of this study that the inquiry course was chosen over other choices seems to be a favorable finding. Attitudinal effects could have been worse in some of the other choices.

It is very important to note that the form of inquiry studied here was termed "guided discovery." Much caution would be needed to generalize the results to other physical science inquiry courses that used inquiry in substantially different ways, perhaps more rigorously or with an historical approach. If personal attention from instructors was an important determinant of attitude, then a more rigorous inquiry course with a similar use of instructors could have similar attitudinal effects; that is, a wish to take more such courses despite a negative attitude shift toward inquiry. Historical or non-laboratory inquiry courses do not have the same instructor situation and the results would not be applicable. For the general student who tends to avoid the rigor of science, a more rigorous course might show a greater decrease in attitude toward inquiry.

When the assumption is recalled that non-science students who enjoy using inquiry might better retain and generalize the ability to use inquiry later, a final conclusion that could be drawn from this study is that students with a physical science inquiry course in their background are not expected to use inquiry later to any more degree than students without that background in inquiry.

Recommendations

If the attitudinal effects of inquiry teaching are to be still more clearly determined, further studies might consider some improvements over the present study. Some recommendations can be inferred from certain internal inconsistencies found in this study.

There was internal evidence of difficulty with the semantic differential as used in this study. Not only was attitude as measured with it found unrelated to attitude as measured behaviorally, but also a strong relation was found between attitude as measured with the semantic differential and student evaluation of instructor, independent of particular instructors. Since the semantic differential for instructor was marked by the students immediately after three semantic differentials for attitude toward the inquiry course, it is plausible that those students who evaluated the inquiry course high went on to evaluate the instructor high. The largest correlation of the study (0.53) was found between posttest attitude and instructor evaluation. The semantic differentials may have interacted, because the argument could also be reversed: those students so favorably predisposed as to rate the instructor high, independently of who he/she was, were also favorably inclined (at posttest time) toward the other semantic differential ratings. Any further study using semantic differentials should attempt to eliminate this possible interaction. Previous studies using semantic differentials on a number of phrases including "instructor" or teacher" did not report an interaction among the semantic differentials.

The fact that students elect to take a second inquiry course despite an unfavorable shift of attitude in the first course showed that attitude toward inquiry as measured by behavior did not agree with attitude toward inquiry as measured on paper. Either two different kinds of attitude were measured, or the paper measure inadequately reflected the behavioral measure. Further studies should investigate further the distinction in attitude and whether the semantic differential can be made to measure what was demonstrated behaviorally.

As the description of the development of the knowledge test indicated, the non-science students in a non-tested inquiry course tended to be a population on which it was very difficult to measure knowledge changes. If the operational definition of cognitive knowledge is taken to be that which, when tested, gives a test with high alpha (greater than 0.90), then substantial improvement is needed in the multiple-choice test method used. Test items on which students prefer the wrong answer on the posttest are a symptom that what was really learned in class was not what was intended, or that there was variation in instructor tolerance of the style of expression used by students. If the course and teaching are taken as fixed, then the test items need rectification. Two test items that appeared on the final version of the test but were not scored, pertained to content nearly every instructor stressed in class, yet few students could answer the items correctly on the posttest. Eliminating such items raises the alpha, yet that does not close the issue, since they supposedly represent valid course content. The question of why some test items on course content are not answered better by students who answer other items better can be resolved only by a much more exhaustive knowledge test that may have to cover several dimensions of knowledge. Such a test would be difficult to develop due to the care required in writing many test items which can be read unambiguously by the non-science student and in making sure that such diverse students do not find more than one item answer acceptable on the posttest.

Since knowledge changed very little in the course studied, the method of measuring knowledge in this form of course needs to be explored more carefully. In a course that is supposed to teach content, a large gain of knowledge is expected. Either the inquiry course for non-science majors was not meant to impart much knowledge, or the students were not learning as much as intended, or the test covered a kind of knowledge different from what was covered in the course. The latter reason is the least likely since no one who participated in the study reported that any aspect of the knowledge test seemed irrelevant. Regardless of whether the course was intended to be a content-oriented course, it did not function effectively as one. In generalizing the results, it must be kept in mind that other inquiry courses could be more content-oriented without sacrifice of inquiry objectives. In other studies on inquiry a congitive knowledge test might be very relevant, and might show much larger gain.

If the negative attitude change did occur and was not a result of measurement problems, then there are at least three commonly discussed reasons why attitude could have changed in that way. As described in Chapter 2, three possible ways attitude could change were through the rewarding of behavior, through cognitive dissonance, or through the internalization of control.

If the course content or subject matter was unrewarding, it was not so unrewarding that the students would not elect another course of similar content. More explicit effects of science subject matter should be investigated with studies of similar inquiry courses in other subjects. Other sources of reward--course methods and instructor--were distinguised by this study. The effect of method apparently was unrewarding whereas the effect of instructor was rewarding.

If cognitive dissoance in the course did not induce favorable attitude changes, it could not have been due primarily to the instructor, for there was clear indication the students favored their interactions with the instructor. Future studies might investigate whether the use of scientific methods diminishes attitudes in this respect, perhaps by there being too much cognitive dissonance despite attempts by instructors to moderate it.

With respect to internalization of control over inquiry processes, it seemed that the instructors might have allowed an appropriate degree of internalization so that the instructors were favorably viewed but that the degree of internalization was insufficient for inquiry to be appreciated. Future studies might compare the amount of internalization needed for mastery of, and appreciation of, inquiry with the amount of control the students want to apportion between themselves and the instructor.

Also valuable but very difficult would be a followup study, preferably on the same sample, as to whether the obtained attitude shift remains stable or becomes later a positive shift. A followup study should also try to ascertain if the students are generalizing the inquiry skills to other aspects of their lives than physical science.

In view of the persistently inconsistent attitudinal results of this and previous studies, further work should proceed much more carefully with what occurs within the inquiry course. Attitude measures should be more detailed and more unobtrusive. The effects of student-student and student-class interactions should be considered. Affective and cognitive learning by inquiry should be studied with a close focus on what the student is doing and feeling as he works with apparatus, with data, with classmates, and with the instructor in the ongoing classroom situation.

More attention should be paid to the effects of science subject matter on attitude. Newtonian mechanics might have been less interesting to the student than, for example, thermal physics or electricity. Other physical science courses have different admixtures of topics with unknown consequences for attitudinal effects.

Appendix A.

Pretest and Posttest Form:

Directions and background questions

Four semantic differentials

Knowledge test

Dear physical science student:

About half an hour of your time is requested for answering the following sets of questions. If you will answer every item as best you can, your assistance will be very valuable in our on-going study of the effectiveness of the physical science courses.

The results will not be seen by your instructor, nor will they be used for the grading of any student in the course.

Take all the time you need so you can try to answer every question. Note that questions are on both sides of the pages.

--John Mauldin, science education graduate student

BACKGROUND

Social security number____________________

Year (circle) freshman sophomore junior senior

Major_______________________________ Sex (circle) female male

College GPA (grade point average)________________________

How many hours of college science are you required to take?______________

Are you required to take six hours of physical science?_________________

How many hours of college science have you completed (do not count PS 303)?_______________

Will you take PS 304? _________________

Which one of these was your principal reason for choosing PS 303 (circle one):

1. recommended by a friend 2. required 3. easy course
4. no lectures 5. sounded interesting 6. (give other)________

Circle one response to each of these questions:

I read about science in newspapers, magazines and books (not course texts):

never rarely sometimes often all the time

I engage in personal activities (not school-required) involving science:

never rarely sometimes often all the time

EVALUATIONS

On the following four pages are a means for you to help evaluate some aspects of this kind of course. At the top of each page is a short phrase. You are asked to react to each phrase by means of the 12 scales given below it. On each scale put an X in one of the 7 blanks to show how much you agree with one of the words as a description of the phrase at the top. The middle blank would be used to show a neutral reaction or if neither word seems to apply. Marking a blank to the left or to the right would show an increasing agreement with the word at the left or the right. Here is an example of two scales applied to the phrase "The world":

FAST :___:___:___:___:___:_X_:___: SLOW
STRONG :_X_:___:___:___:___:___:___: WEAK

The positions of the X here show moderate agreement that the world is "slow," and strong agreement that the world is "strong" rather than "weak."

PHYSICAL SCIENCE

GOOD	:____	:____	:____	:____	:____	:____	:____:	BAD
SMALL	:____	:____	:____	:____	:____	:____	:____:	LARGE
FAST	:____	:____	:____	:____	:____	:____	:____:	SLOW
UNPLEASANT	:____	:____	:____	:____	:____	:____	:____:	PLEASANT
STRONG	:____	:____	:____	:____	:____	:____	:____:	WEAK
QUIET	:____	:____	:____	:____	:____	:____	:____:	ACTIVE
CLEAN	:____	:____	:____	:____	:____	:____	:____:	DIRTY
LIGHT	:____	:____	:____	:____	:____	:____	:____:	HEAVY
HOT	:____	:____	:____	:____	:____	:____	:____:	COLD
WORTHLESS	:____	:____	:____	:____	:____	:____	:____:	VALUABLE
SOFT	:____	:____	:____	:____	:____	:____	:____:	HARD
DULL	:____	:____	:____	:____	:____	:____	:____:	SHARP

DOING EXPERIMENTS

GOOD	:____:____:____:____:____:____:____:	BAD
SMALL	:____:____:____:____:____:____:____:	LARGE
FAST	:____:____:____:____:____:____:____:	SLOW
UNPLEASANT	:____:____:____:____:____:____:____:	PLEASANT
STRONG	:____:____:____:____:____:____:____:	WEAK
QUIET	:____:____:____:____:____:____:____:	ACTIVE
CLEAN	:____:____:____:____:____:____:____:	DIRTY
LIGHT	:____:____:____:____:____:____:____:	HEAVY
HOT	:____:____:____:____:____:____:____:	COLD
WORTHLESS	:____:____:____:____:____:____:____:	VALUABLE
SOFT	:____:____:____:____:____:____:____:	HARD
DULL	:____:____:____:____:____:____:____:	SHARP

MAKING INFERENCES FROM OBSERVATIONS

GOOD :____:____:____:____:____:____:____: BAD

SMALL :____:____:____:____:____:____:____: LARGE

FAST :____:____:____:____:____:____:____: SLOW

UNPLEASANT :____:____:____:____:____:____:____: PLEASANT

STRONG :____:____:____:____:____:____:____: WEAK

QUIET :____:____:____:____:____:____:____: ACTIVE

CLEAN :____:____:____:____:____:____:____: DIRTY

LIGHT :____:____:____:____:____:____:____: HEAVY

HOT :____:____:____:____:____:____:____: COLD

WORTHLESS :____:____:____:____:____:____:____: VALUABLE

SOFT :____:____:____:____:____:____:____: HARD

DULL :____:____:____:____:____:____:____: SHARP

PHYSICAL SCIENCE INSTRUCTOR

GOOD	:____:____:____:____:____:____:____:	BAD
SMALL	:____:____:____:____:____:____:____:	LARGE
FAST	:____:____:____:____:____:____:____:	SLOW
UNPLEASANT	:____:____:____:____:____:____:____:	PLEASANT
STRONG	:____:____:____:____:____:____:____:	WEAK
QUIET	:____:____:____:____:____:____:____:	ACTIVE
CLEAN	:____:____:____:____:____:____:____:	DIRTY
LIGHT	:____:____:____:____:____:____:____:	HEAVY
HOT	:____:____:____:____:____:____:____:	COLD
WORTHLESS	:____:____:____:____:____:____:____:	VALUABLE
SOFT	:____:____:____:____:____:____:____:	HARD
DULL	:____:____:____:____:____:____:____:	SHARP

PHYSICAL SCIENCE KNOWLEDGE

An answer sheet is supplied for this part. First turn it to locate the corner where "student number" is to be put and write your social security number there.

Then mark your social security number by blackening one digit per column, using the soft pencil supplied. As you probably know, you should just fill the outlined slot with pencil, neither too much nor too little, and make no other stray marks.

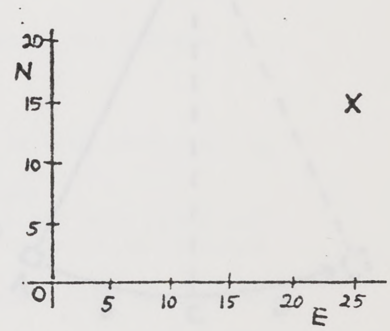
Next, to the left, mark the unique number of this section in the first five (unlabeled) columns. Leave column six blank.

Now answer each of the following 23 questions. Choose the best or correct answer and darken its letter (A, B, C, D, or sometimes E) on the answer sheet under the appropriate number 1 thru 23. The questions are on both sides of the pages.

Please do not write on these pages. If you need scratch paper, use the back of the "Background" page.

1. A good timing device should
   A. produce a steady ticking sound
   B. be easily carried about from place to place
   C. have some sort of regularly repeating action
   D. be inexpensive

2. On finding that a certain object occupies as much space as 10 grams of water, it can be said that the volume of the object is approximately
   A. 10 cm
   B. 10 cm^2
   C. 10 cm^3
   D. 10 ml^2
   E. 10 ml^3



3. In the diagram shown, an object is moved 40 units from the origin (0) to the point X. Which one of these statements is true?
   A. the displacement of the object from 0 is 25 units
   B. the distance of the object from 0 is 25 units
   C. the object's displacement is on a 45° line
   D. the object is 25 units east, 15 units north of 0
   E. the object is 25 units north, 15 units east of 0

4. Magnetic force and gravitational force are alike in that both
   A. vary in direct proportion to distance
   B. decrease as distance between the objects producing the force increases
   C. are capable of not only attracting but also repelling objects
   D. attract only certain kinds of objects
   E. act only at close distance

5. If a parson walks ten blocks east, five blocks north, three blocks west, and five blocks south, his displacement from the original point is
   A. three blocks east
   B. three blocks west
   C. seven blocks east
   D. thirteen blocks east
   E. twenty three blocks northeast

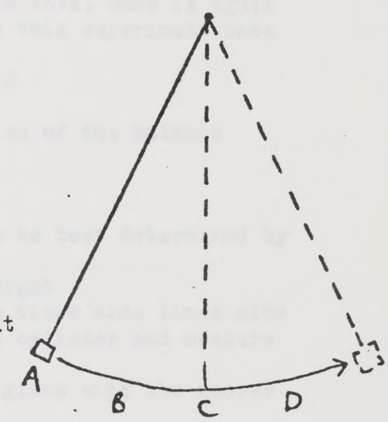
TURN TO THE OTHER SIDE →

Continue marking the best or correct answer on the answer sheet:

6. A pencil and a battery fell from a laboratory desk and struck a student's foot. The impact of the pencil did not hurt, but the battery bruised his foot. The best explanation in terms of physical science is
   - A. the mass of the battery was greater than that of the pencil
   - B. the acceleration and velocity of the battery were greater than that of the pencil
   - C. the mass and acceleration of the battery were greater than that of the pencil
   - D. the speed of the battery was greater than that of the pencil
   - E. the acceleration of the battery was greater than that of the pencil

7. A very important difference between weight and mass is
   - A. weight is affected by volume and mass is not
   - B. weight varies with location and mass is defined as being constant under a variety of conditions
   - C. mass is determined by size, and weight by density
   - D. weight is the amount of matter in an object and mass is a measure of gravitational attraction
   - E. mass is affected by volume and density;weight is not

The pendulum shown below is released at point A and allowed to swing. Answer the next four questions for this situation:



8. The speed of the weight is minimum at point
   - A.
   - B.
   - C.
   - D.

9. The speed of the weight is maximum at point
   - A.
   - B.
   - C.
   - D.

10. The acceleration of the weight is minimum at point
    - A.
    - B.
    - C.
    - D.

11. The acceleration of the weight is maximum at point
    - A.
    - B.
    - C.
    - D.

12. To be useful a standard of measurement must be
    - A. clearly defined and accurately reproducible
    - B. expressed in the metric system of units
    - C. recognized as being without error
    - D. based on accepted values for the size of the foot and pound
    - E. used by the majority of the people in the world

13. By careful measurement it was determined that a certain substance had a volume of 16 cm^3 and a mass of 4 grams. Its density is

A. 4 gm/cm^3
B. 0.25 gm/cm^3
C. 16 cm^3/gm
D. 4 cm^3/gm
E. 0.25 cm^3/gm

14. Hilda was given a sealed box to examine. She was asked to make some observations about the box. After shaking it, tilting it, feeling its weight, and listening to the sounds it made, she wrote down the four following statements. The box was sealed and was not opened. Which one of the following could best be called an observation?

A. "There are two objects in the box."
B. "I heard a large object roll and a small one slide when I tilted the box."
C. "I felt two objects fall, one large and one small, when I tilted the box."
D. "I heard two different sounds and felt two different vibrations when I tilted the box."
E. All of the above could be called observations.

15. In an experiment using the equal arm balance, you measure separately the masses of some sodium bicarbonate, a container, and some acetic acid. The total mass is 145.54 grams. You then drop the sodium bicarbonate into the acetic acid and a vigorous fizzing occurs. After the action has stopped, the total mass is again determined and found to be 143.62 grams. From having done this experiment once you can safely conclude

A. that the law of conservation of matter is not valid
B. that the mass is conserved
c. that the mass remained the same within the precision of the balance
D. that the mass of the contents decreased
E. none of the above

16. The volume of an irregularly shaped piece of glass can be best determined by the following process:

A. measure it and multiply length times width times height
B. form it into a cube, measure it, and multiply side times side times side
C. place it in a known volume of water in a graduated cylinder and measure the increase in water level
D. use a balance to compare the mass of the piece of glass with the masses of various known volumes of glass
E. there is no way to determine the volume without altering the substance

17. Area can best be defined as

A. length times width
B. the space taken up by matter
C. displacement times displacement
D. the interval between two lines
E. the extent of surfaces

18. Time is best defined as

A. the interval between two events
B. the natural division of the day into hours, minutes, and seconds
C. the measurement of change
D. that aspect of a system that does not change

TURN TO THE OTHER SIDE →

19. To calculate the speed of an object it is necessary to know its
   A. mass and time of travel
   B. mass and distance traveled
   C. time and distance traveled
   D. distance traveled and rate of acceleration
   E. where it started

20. In which one of the following is matter <u>not</u> conserved?
   A. boiling a liquid
   B. crushing a rock
   C. any chemical reaction
   D. turning on a light
   E. none of these

21. To find the effect of the length of the string on the rate at which a pendulum swings, one would have to
   A. hold the string length constant and vary the number of swings timed
   B. hold the string length constant and vary the weight on the end of the string
   C. vary the string length and hold the weight constant
   D. remove the weight and time the pendulum
   E. measure the speed with which the weight moves

22. A meter stick is
   A. more accurate than a yardstick
   B. longer than a yardstick
   C. shorter than a yardstick
   D. used only in science laboratories
   E. calibrated based on the standard foot

23. The width of a room could, with care, be measured to an accuracy of
   A. $\pm$ 1 ft
   B. $\pm$ 1 cm
   C. $\pm$ 1 inch
   D. $\pm$ 1 mm
   E. $\pm$ 1 meter

Please check that you marked correctly your social security number on the answer sheet. Without it we cannot use your participation in this study because your other responses and your answer sheet will become separated.

Appendix B. Letter to instructors

3 Sept. 1974

Dear physical science instructor:

I am asking for your voluntary assistance in my study of the effectiveness of our physical science courses, a study I am carrying out for my doctorate under the supervision of Dr. Little. You are asked to administer the giving of a set of written questions at the very beginning of your first class, and again at the end of the course. The students will require no more than half an hour to answer the questions as they work at their own paces.

I will deliver the materials to you in a packet after receiving your consent, sometime before your first class on Wed., 4 Sept. or Thurs., 5 Sept. As you hand out the materials, you need not make any explanation of what is to happen or why, beyond that given in the brief introduction which the students will read. The general atmosphere is that you are allowing me to conduct this study and assisting in administering it.

When handing out the materials, observe the following steps:

. Put the unique number of your section on the board.

. They are to work on the "Background and evaluation" pages first. They write directly on those pages.

. Later they start on the "Knowledge" pages and must use the answer sheets and soft pencils. They should not write anything on the knowledge question sheets; just mark the answer sheets.

. When it seems like nearly everyone is finished, collect all materials. These can be collected in one bundle if care is taken not to wrinkle the answer sheets. I would prefer that the "Knowledge" question sheets be in a separate pile if this is convenient.

. Place everything except the pencils back in the envelope provided and close it.

To anticipate several questions they might ask, I offer these remarks:

. They are not to write their names on anything.

. Their social security number is needed on both the "Background" sheet and on the answer sheet. It will be used only by me, and then only so that I can keep track of the two parts and be able to check back later with the same students for follow-up.

. Neither set of questions can appropriately be called a "test" and should not be so called. Nor is evaluation of the instructor or his section an objective of my study.

I will return to collect the materials. If I do not, you may give them to Dr. Little or slide them under my office door, room 8.308.

Thanks from:

John Mauldin
Science education doctoral candidate

Appendix C: Analysis of knowledge test

Item	Correct answer	Pretest: Percent correct	Pretest: Correlation with total	Pretest: Percent Answering: A	B	C	D	E	Posttest: Percent correct	Posttest: Correlation with total	Posttest: Percent Answering: A	B	C	D	E
1	C	81	.25	12	6	81	1	*	91	.27	6	2	91	1	*
2	C	30	.41	22	22	30	15	9	58	.38	7	16	58	7	11
3	D	80	.34	5	5	8	80	1	91	.28	2	1	3	91	3
4	B	53	.29	21	53	18	5	3	50	.27	34	50	14	1	1
5	C	85	.31	2	7	85	2	3	84	.35	4	4	84	3	5
6	A	51	.38	51	6	40	1	3	61	.43	61	9	29	1	1
7	B	27	.32	4	27	37	10	21	42	.38	4	42	26	12	15
8	A	42	.37	42	5	11	41	*	56	.42	56	4	12	28	*
9	C	52	.44	10	17	52	21	*	53	.47	10	14	53	23	*
10#	C	16	#	18	13	16	52	*	9	#	29	14	9	48	*
11#	A	24	#	24	41	20	15	*	13	#	13	38	30	19	*
12	A	77	.35	77	5	6	2	10	83	.31	83	3	4	2	8
13	B	24	.38	19	24	15	29	6	26	.42	21	26	14	31	7
14	D	30	.24	5	4	3	30	58	47	.38	5	7	1	47	41
15	D	54	.19	2	6	9	54	27	54	.25	3	7	8	54	28
16	C	64	.44	8	3	64	10	14	78	.48	7	4	78	6	4
17	E	18	.44	56	24	1	1	18	22	.40	52	22	2	2	22
18	A	41	.37	41	21	31	6	*	75	.44	75	6	15	4	*
19	C	58	.42	5	13	58	24	0	71	.38	3	6	71	19	0
20	E	40	.24	26	4	11	18	40	46	.33	27	6	10	12	46
21	C	65	.41	6	10	65	4	13	75	.49	8	10	75	2	4
22	B	56	.43	23	56	15	1	5	70	.42	21	70	3	2	4
23	D	52	.31	5	21	13	52	8	61	.25	2	25	7	61	4

Notes:

--pretest: alpha=0.64, N=300, mean score=10.82, sigma=3.39

--posttest: alpha=0.68, N=300, mean score=12.94, sigma=3.48

--correlation is the correlation coefficient between item answers and total score on 21-item test.

--# denotes items not included in scoring 21-item test.

--* denotes items where choice E is not available.

Appendix D. Glossary of some terms used in educational research

Analysis of variance--When subjects are treated in two or more groups, part of the variance within the groups can be attributed to the difference in treatment. In the analysis, the variance of all groups of subjects taken together is compared to the variance of the several group means, and an F-test for significance is performed.

Correlation coefficient--When the scores x_i and y_i are obtained from each of N subjects, the correlation coefficient of the scores is defined as:

$$r \equiv \frac{\sum x_i y_i - NM_x M_y}{N \sigma_x \sigma_y}$$

where M_x is the mean of the x-scores and M_y is the mean of the Y-scores, and σ_x and σ_y are the respective standard deviations of the scores.

Degrees of freedom--The number of ways in which a number computed from scores can vary. If the mean score for N subjects is computed, the mean has N degrees of freedom, and N-1 independent degrees of freedom. Once N-1 scores are fixed and the mean is known, the remaining score is completely determinant.

F-test--A test for statistical significance applied to the F-ratio or ratio of the variance within to the variance between groups. Tables are available with which it can be determined whether a given F, with degrees of freedom determined by the number of subjects and by the number of groups, exceeded chance level with a probability of 0.05 or 0.01.

Likert scale--Generally, a five-position scale with "agree" at one end, neutral in the middle, and "disagree" at the other end, to be applied to a statement about which the subject's opinion or attitude is sought.

Standard deviation--The root-mean-squared deviations of scores from their mean, usually assuming a normal or Gaussian distribution of scores, also called sigma.

t-test--A test for statistical significance applied to t, which can be calculated by

$$t = \frac{M_x - M_y}{S}$$

where M_x and M_y are the means of x-scores and y-scores respectively, and S is the standard error. S is estimated from the sigmas of the scores by various methods. The t-test tests whether two sample means had a significant difference.

Variance--The square of sigma, or the mean-squared deviations of scores from their mean.

Appendix E. Introduction to the student for Physical Science 303/304

TO THE STUDENT

This course assumes that you, as a student beginning a college program, have had a minimal previous training in science and mathematics. It attempts to reinforce your ability to observe phenomena, to infer that certain concepts are useful in describing phenomena, to perceive that relationships exist connecting the concepts as natural laws, to construct abstract models to explain and relate phenomena, to extrapolate from primitive concepts other compound or secondary concepts, and to elaborate theories as logical systems of concepts and laws to explain and predict phenomena of the physical world. It attempts to do this by having you perform these actions for a sequence of phenomena which has been carefully selected from the infinite possibilities among natural physical events. People look at things in different ways and what you might initially consider as a useful concept possibly would not be considered as useful by some one else. A concept, law or theory is useful only if it is generally accepted as explaining and predicting phenomena. The criterion for establishment of a concept, law or theory is thus nothing more than general concurrence in its value. You, and the rest of the class, will decide in class discussion what are the useful concepts and laws which can be inferred from your observations.

Most of the concepts you will study in this course come from phenomena you could meet in ordinary experience and whose names may be known to you. Nevertheless few people understand these concepts and the laws which connect them well enough to make effective use of what knowledge they have. In addition to a deeper understanding of the basic concepts of physical science, your studies in this course should improve your ability to observe and analyze events in your other activities. Physical science is a useful field for practicing these techniques since it contains the least ambiguous concepts and laws of any field which treats of actual events. Further, events in physical science are repeatable whereas they are not in history, economics or in most fields of human endeavor. This repeatability permits more detailed study under more easily specified and controlled conditions.

The pages which follow give the bare minimum of directions necessary to produce the phenomena to be studied. Your first trial may make you aware of other conditions which need to be controlled so that you will need to start over with a modified procedure. You will keep a record book of your observations (data), assumptions and inferences as well as those mutually agreed on by the class, your scientific community. The direction sheets from this book plus your record book will serve as your text book for the course. Enter all ideas, doubts, questions, etc. directly into your record book - NOT on loose sheets for recopying. Cross out errors instead of erasing.

Your grade in the course is based primarily on the instructor's opinion of your performance in class. If examinations are given they are to help you clarify your ideas and not primarily for grade evaluation. Your instructor will use your record book, his talks with you in class, possibly examination grades, and your contributions to class discussion in his evaluation of your performance. The course is intended to help all of you master the material and not to present a series of barriers which only some of you can overcome.

Attendance is much more important than in a lecture type course since your performance is done almost completely in the class time. In addition, the material is highly sequential with each concept and law being used repeatedly in the development of further concepts and laws.

Appendix F. Alpha reliability

The alpha reliability of a test is defined by the formula

$$\alpha \equiv \frac{N}{N-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_T^2}\right)$$

where N is the number of items on the test, σ_i is the variance of the distribution for the item i, and σ_T is the variance of the total score. This formula, from Cronbach (1951), assumes that the answers to a test item fall on an ordered scale, which is not the case for a multiple-choice test unless the items are scored right=1 and wrong=0 .

Cronbach has shown the following theorems about alpha reliability (assuming all scores are normally distributed):

(1) Alpha is the mean of all possible split-half reliabilities, where a split-half reliability is the correlation of the total score on one-half the test with the total score on the other half.

(2) Alpha is the correlation between scores on randomly sampled subsets of the test items.

(3) Alpha is less than or equal to the proportion of test variance due to commonality among the test items.

The Scholastic Aptitude Test (SAT) has an alpha of 0.90. The Graduate Record Examination (GRE) has alpha=0.91 on the verbal part and 0.90 on the quantitative. The GRE Advanced Test in physics has alpha=0.90 (Buros, 1972). The College Board Advanced Placement test in physics has alpha=0.65 on the objective questions (Buros, 1965). Experience is the only guide to what constitutes an acceptable alpha. These quoted values illustrate the values obtained on large-scale well-known tests.

An alpha of 0.7 to 0.8 is quite good for an attitude measure such as the semantic differential. An alpha greater than 0.9 is generally expected for a test of cognitive knowledge, where it is desired that test items consistently sum to a reliable score.

For the special case of a test scored as right-wrong, alpha reduces to the Kuder-Richardson formula 20:

$$\alpha = \frac{N}{N-1}\left(1 - \frac{\sum p_i q_i}{\sigma_T^2}\right) ,$$

where p_i is the proportion (between 0 and 1) passing test item i, and q_i is the proportion failing that item.

Alpha is large if the item scores have compact distributions (small σ_i^2) compared to the total test score distribution (with variance σ_T^2). The alpha reflects the consistency with which a total score represents an appropriate fraction of test items answered correctly. The extreme cases of a zero total score or the maximum total score both have alpha=1, since in such cases all test items were answered perfectly consistently.

References

Bruner, Jerome S. On knowing: essays for the left hand, 1962, Atheneum.

Bruner, Jerome S. The process of education, 1960, Random House.

Buros, Oscar. Sixth Mental measurements yearbook, 1965, Gryphon Press.

Buros, Oscar. Seventh Mental measurements yearbook, 1972, Gryphon Press.

Cohen, Robert S. "Individuality and common purpose: the philosophy of science," Science Teacher, 31, pp. 27-33, May 1964.

Commoner, Barry. "Is science getting out of hand?" Science Teacher, 30, pp. 11-16, Oct. 1963.

Cronbach, Lee J. "Coefficient alpha and the internal structure of tests," Psychometrika, 16, pp. 297-334, 1951.

Diehl, Tennieson H.T. A study of attitude and thought pattern changes resulting from the use of a physical science course for nonscience majors. Unpublished dissertation, Ohio State University, 1967.

Education Commission of the States. "Science Objectives for 1972-73 assessment," 1972.

Fellers, William Oscar. The change in attitude toward science upon completion of a one semester general education physical science course at the junior college level. Unpublished dissertation, University of Northern Colorado, 1972.

Festinger, Leon. A theory of cognitive dissonance, 1957, Stanford Univ. Press.

Frangos, George John. A study of three outcomes of a college level course in physical science for non-science students (adapted PSNS). Unpublished dissertation, Ohio State University, 1971.

Gagne, Robert. "Learning requirements for enquiry," J.Res.Sc.Teaching, 1, pp. 144-153, 1963.

Grozier, Joseph Edward. The role of the laboratory in the development of positive attitudes toward science in a college general education science course for nonscientists. Unpublished dissertation, East Texas State University, 1969.

Guilford, J. P. Fundamental statistics in psychology and education, 1973, McGraw-Hill.

Gunsch, Leonhardt Maurice. A comparison of students' achievement and attitude changes resulting from a laboratory and nonlaboratory approach to general education physical science courses. Unpublished dissertation, University of Northern Colorado, 1972.

Hecht, Alfred Roland. A semantic differential evaluation of attitudinal outcomes of introductory physical science. Unpublished dissertation, University of Illinois (Urbana), 1970.

Himaya, Makram I. Identification of possible variables for predicting student changes in physical science courses designed for nonscience majors. Unpublished dissertation, University of Iowa, 1972.

Jones, Haskell Lee. The development of certain aspects of scientific literacy in college freshman physical science students. Unpublished dissertation, Ohio State University, 1969.

Kerlinger, Fred N. and Pedhazur, Elazar J. Multiple regression in behavioral research, 1973, Holt-Rinehart Winston.

Koelsche, C. L. and Morgan, A. G. Scientific literacy in the sixties, 1964, University of Georgia.

Krathwohl, David R.; Bloom, Benjamin S.; and Masia, Bertran B. Taxonomy of educational objectives: affective domain, 1964, David McKay Co.

Little, Robert N. Introductory college physical science, 1973, unpublished.

Little, Robert N. Instructor's guide for introductory college physical science, 1973, unpublished.

Little, Robert N. "Preservice formation of physics teachers: technical education." Paper presented at International Congress on the Education of Teachers of Physics in Secondary Schools, 11-17 Sept. 1970, Eger, Hungary.

McREL (Mid-Continent Regional Education Laboratory). Inquiry objectives in the teaching of biology, 1969, published by McREL, Kansas City, Missouri.

Milson, James L. "The development and evaluation of physical science curriculum materials designed to improve student attitudes," J.Research in Science Teaching, 9, pp. 289-304, 1972.

Osgood, Charles; Suci, George; and Tannenbaum, Percy. The measurement of meaning, 1957, University of Illinois Press.

Oshima, Eugene Akio. Changes in attitude toward science and confidence in teaching science of prospective elementary teachers. Unpublished dissertation, Oklahoma State University, 1966.

Rokeach, Milton. Beliefs, attitudes, values, 1968, Jossey Bass Pub. Co.

Rotter, J. B. "Generalized expectancies for internal versus external control of reinforcement," Psych. Monographs, 80, no. 609, 1966.

Rutherford, F. James. "Role of inquiry in science teaching," J.Res.Sc. Teaching, 2, pp. 80-84, 1964.

Schulman, Lee. "Psychological controversies in the teaching of science and mathematics," Science Teacher, Sept. 1968, p. 34.

Spradlin, Susan Dunbar. A study of change in science teaching behavior in the classroom following involvement in institute training or retraining programs. Unpublished dissertation, University of Texas at Austin, 1973.

Steiner, Harrel. A study of the relationships between teacher practices and student performances of selected inquiry process behaviors in the affective domain in high school biology classes. Unpublished dissertation, University of Texas at Austin, 1970.

Stekel, Frank D. The effects of increasing student involvement in a college physical science laboratory program. Unpublished dissertation, Indiana University, 1970.

Webb, Eugene J.; Campbell, Donald T.; Schwartz, Richard; and Sechrest, Lee. Unobtrusive measures: nonreactive research in the social sciences, 1966, Rand McNally.

Wade, Nicholas. "Science still in public favor," Science, 182, p. 369, 26 Oct. 1973.

Wernegreen, Johannes Oscar. An exploratory study of changes in concepts as measured by a semantic differential instrument during a college physical science course for nonscience students. Unpublished dissertation, University of Illinois (Urbana), 1971.

Wilkerson, George J. Changing student attitudes, 1972, private publication.

Wood, Elizabeth A., "Physical science for nonscience students," CCP Newsletter, no. 17, Oct. 1968, pp. 5-8.

Wood, Elizabeth A., "PSNS project at RPI," Amer. J. Physics, 34, no. 9, part 2, Sept. 1966, pp. 891-894.

Veldman, Donald J. "COMMAP: A computer program for commonality analysis," unpublished paper, Project PRIME, University of Texas at Austin, 1975.

Veldman, Donald J. Edstat-V user manual, 1971, ibid.

Veldman, Donald J. Fortran programming for the behavioral sciences, 1967, Holt-Rinehart Winston.

_______, "Report on Commission on College Physics," Amer. J. Physics, 34, no. 9, part 2, Sept. 1966.

The vita has been removed from the digitized version of this document.